LE SCIENZE
edizione italiana di SCIENTIFIC AMERICAN

numero 198
febbraio 1985
anno XVIII
volume XXXIV

La memoria degli atomi

Sistemi atomici decaduti da stati ordinati possono essere indotti a recuperare l'ordine iniziale, il cui grado di ripristino consente di realizzare lo studio di alcune interazioni che risultano di norma difficilmente osservabili

di Richard G. Brewer ed Erwin L. Hahn

Nel 1872 Ludwig Boltzmann, uno dei fondatori della termodinamica moderna, tenne una conferenza durante la quale affermò che l'entropia, o disordine, dei sistemi isolati aumenta irreversibilmente nel tempo. Nell'udire questa affermazione il fisico Joseph Loschmidt si alzò a protestare. Egli sosteneva che le leggi che governano il moto di tutte le particelle sono simmetriche rispetto al tempo, quindi qualunque sistema che fosse decaduto dall'ordine al caos si poteva far ritornare in uno stato ordinato semplicemente invertendo la quantità di moto di ogni particella, senza modificare l'energia cinetica totale del sistema. Con aria di disprezzo Boltzmann puntò l'indice verso Loschmidt e disse: «*Tu* inverti le quantità di moto.»

Questo scontro accademico illustra la natura paradossale del secondo principio della termodinamica, secondo il quale i sistemi tendono verso la massima entropia. Eppure le idee di Loschmidt restano irrefutabili. Se potessimo filmare il moto di un qualsiasi gruppetto di particelle e mostrare il film a un fisico, egli non potrebbe in alcun modo dirci se il proiettore stava girando in avanti o all'indietro. Di conseguenza, secondo l'interpretazione di Loschmidt (che doveva essere chiamata paradosso di Loschmidt), qualsiasi legge che regola il comportamento di grandi raggruppamenti di particelle dovrebbe essere simmetrica rispetto al tempo. Mentre il significato e le conseguenze del secondo principio sono ancora vivi argomenti di ricerca e di discussione (si veda *La freccia del tempo* di David Layzer in «Le Scienze», n. 92, aprile 1976), esistono oggi numerosi metodi con i quali realizzare l'inversione temporale di Loschmidt. In altre parole, un sistema di particelle decaduto da uno stato altamente ordinato può essere fatto tornare in tale stato invertendo le direzioni di moto (o qualche altro grado di libertà) delle sue particelle costituenti. In realtà un insieme di atomi può mostrare una memoria delle sue condizioni precedenti.

Perché un sistema possa presentare questo tipo di memoria atomica, lo si deve preparare in modo che abbia un certo tipo di ordine, spesso nascosto, nel suo stato apparentemente disordinato. Nei sistemi atomici dei quali ci occuperemo, questo ordine nascosto si ottiene esponendo i campioni (che possono essere solidi, liquidi o gassosi) a radiazione elettromagnetica coerente di vari tipi, tra cui radioonde, microonde e fasci laser. Anche le onde sonore sono in grado di svolgere questa mansione. La ricomparsa di uno stato ordinato in tali sistemi diventa evidente quando il campione emette il proprio impulso elettromagnetico coerente, un'eco della precedente radiazione. A parte il loro interesse intrinseco, questi impulsi di eco e relative forme di radiazione coerente forniscono nuovi metodi di studio del comportamento fondamentale delle interazioni atomiche.

Si può illustrare il concetto di ordine nascosto ricorrendo a un'analogia. Si consideri un gruppo di corridori allineati sulla linea di partenza di una pista circolare *(si veda l'illustrazione a pagina 15)*. Quando lo starter dà il via alla gara con un colpo di pistola, i corridori si distribuiscono lungo la pista correndo ciascuno con una velocità differente dall'altro. Dopo aver percorso la pista diverse volte, alcuni corridori ne avranno superato altri e non vi sarà alcuna evidente correlazione tra le posizioni relative dei corridori e le loro varie velocità. Chi non avesse assistito alla partenza della corsa potrebbe pensare che non vi sia stato alcun ordine particolare nella disposizione dei corridori e che essi costituiscano un sistema disordinato.

Supponiamo ora che i corridori siano costretti a invertire il senso di marcia e a ripercorrere a ritroso la pista a un segnale prestabilito (per esempio, un secondo colpo dello starter) dato dopo t minuti di corsa. Se tutti i corridori mantengono la loro velocità, anche se in senso inverso, essi si ritroveranno ad attraversare simultaneamente la linea di partenza esattamente dopo $2t$ minuti dall'inizio della corsa. In tal modo avranno ripristinato il loro ordine iniziale. (Ordine che scomparirà ancora ovviamente una volta superata la linea di partenza.)

Non si deve ignorare il caso ancora più semplice in cui tutti i corridori procedono con la stessa velocità periferica. In questo caso l'ordine iniziale viene mantenuto e non c'è alcun bisogno di far invertire il senso di marcia. Questo esempio ha un'analogo elettromagnetico, chiamato effetto di «decadimento per induzione libera», che ha trovato numerose applicazioni nelle regioni di frequenza radio e ottica.

Si può illustrare un esempio più concreto di effetto di memoria con metodi meccanici. Si ponga un liquido viscoso nello spazio anulare compreso tra due cilindri di plastica concentrici. Mentre il cilindro esterno è fermo, quello interno è libero di ruotare attorno al proprio asse. Un filetto di liquido colorato, che rappresenta un allineamento iniziale di particelle, viene

iniettato nel liquido. Quando si fa ruotare il cilindro interno, il colorante diffonde in tutto il volume. Se mostrassimo questo volume tra i due cilindri a un esperto di termodinamica, egli direbbe che il colorante è in uno stato di completo disordine (cioè che l'entropia ha un valore massimo) e che il processo di mescolamento è completo e irreversibile. In realtà, il liquido si trova in uno stato di ordine nascosto (o di entropia costante): l'inversione del senso di rotazione del cilindro interno fa invertire il processo di mescolamento e dopo un ugual numero di giri in senso inverso ricompare il filetto colorato.

Nel 1950 uno di noi (Hahn), allora all'Università dell'Illinois, scoprì un effetto di memoria simile in linea di principio ai casi dei corridori e del filetto colorato, ma operante su scala atomica. Un campione di glicerina venne immerso in un campo magnetico e sottoposto a due brevi impulsi di radiazione elettromagnetica a radiofrequenza (rf), separati da un intervallo di tempo t di qualche centesimo di secondo. Il campione mantenne in memoria la sequenza di impulsi e dopo $2t$ secondi dal primo impulso rf emetteva a sua volta un terzo impulso, un'eco. Questo fenomeno è noto come eco di spin nucleare.

L'eco di spin nucleare è una conseguenza delle proprietà giromagnetiche di nuclei atomici, qual è il protone che costituisce il nucleo della maggior parte degli atomi di idrogeno. Dal momento che il protone possiede uno spin e carica elettrica, esso possiede un momento magnetico per molti aspetti simile al momento angolare di un giroscopio. L'asse di spin di un protone disallineato con un campo magnetico costante, come l'asse di un giroscopio posto in rotazione in un campo gravitazionale costante, precede, cioè descrive una circonferenza attorno a una direzione parallela al campo di forze *(si veda l'illustrazione a pagina 16)*. La frequenza di precessione, velocità con cui l'asse del protone descrive la circonferenza, dipende in parte dall'intensità del campo magnetico esterno. Questa tendenza dell'asse di spin del protone a precedere attorno a un campo applicato costante è la base dell'effetto di eco di spin.

In un esperimento di eco di spin le componenti costanti degli assi di spin dei protoni del campione sono all'inizio allineate parallelamente a un campo magnetico esterno costante. Essendo perfettamente parallele al campo, esse non precedono (come non subirebbe precessione un giroscopio perfettamente verticale). A questo punto si applica il primo impulso a radiofrequenza. Questo impulso rf contiene una componente polarizzata circolarmente - un piccolo campo magnetico rotante - che ruota con la stessa velocità con la quale precederebbero gli assi di spin dei protoni se fossero disallineati con il campo costante e se questo fosse il solo campo presente.

L'impulso rf fa eseguire all'insieme di protoni un moto complicato, descritto meglio come la combinazione di due moti di precessione meno complessi *(si veda l'illustrazione a pagina 18)*. Il più semplice di questi due moti è una precessione attorno al campo statico esterno: l'impulso rf disallinea gli assi di spin rispetto al campo costante (anche se gli assi rimangono praticamente allineati l'uno rispetto all'altro durante la breve durata dell'impulso) ed essi cominciano a precedere attorno alle linee di forza del campo statico con la loro normale velocità di precessione. Dato che anche il campo magnetico presente nell'impulso ruota con la stessa velocità, l'angolo tra il campo rf rotante e qualsiasi asse di spin dei protoni si mantiene costante durante la precessione degli spin dei protoni. Dal punto di vista dei protoni il campo impulsato appare di direzione costante. Di conseguenza gli spin dei protoni precedono sia attorno al campo impulsato sia attorno al campo costante. La combinazione di queste due precessioni è una spirale verso il basso descritta dagli assi di spin di ogni protone.

L'angolo di cui si inclinano i protoni è determinato dall'intensità e dalla durata dell'impulso rf. In un tipico esperimento di eco di spin questo primo impulso potrebbe avere una durata tale da inclinare gli spin dei protoni esattamente di 90 gradi dalla verticale; in altri termini, essi vengono a trovarsi sul piano perpendicolare alla loro orientazione originaria.

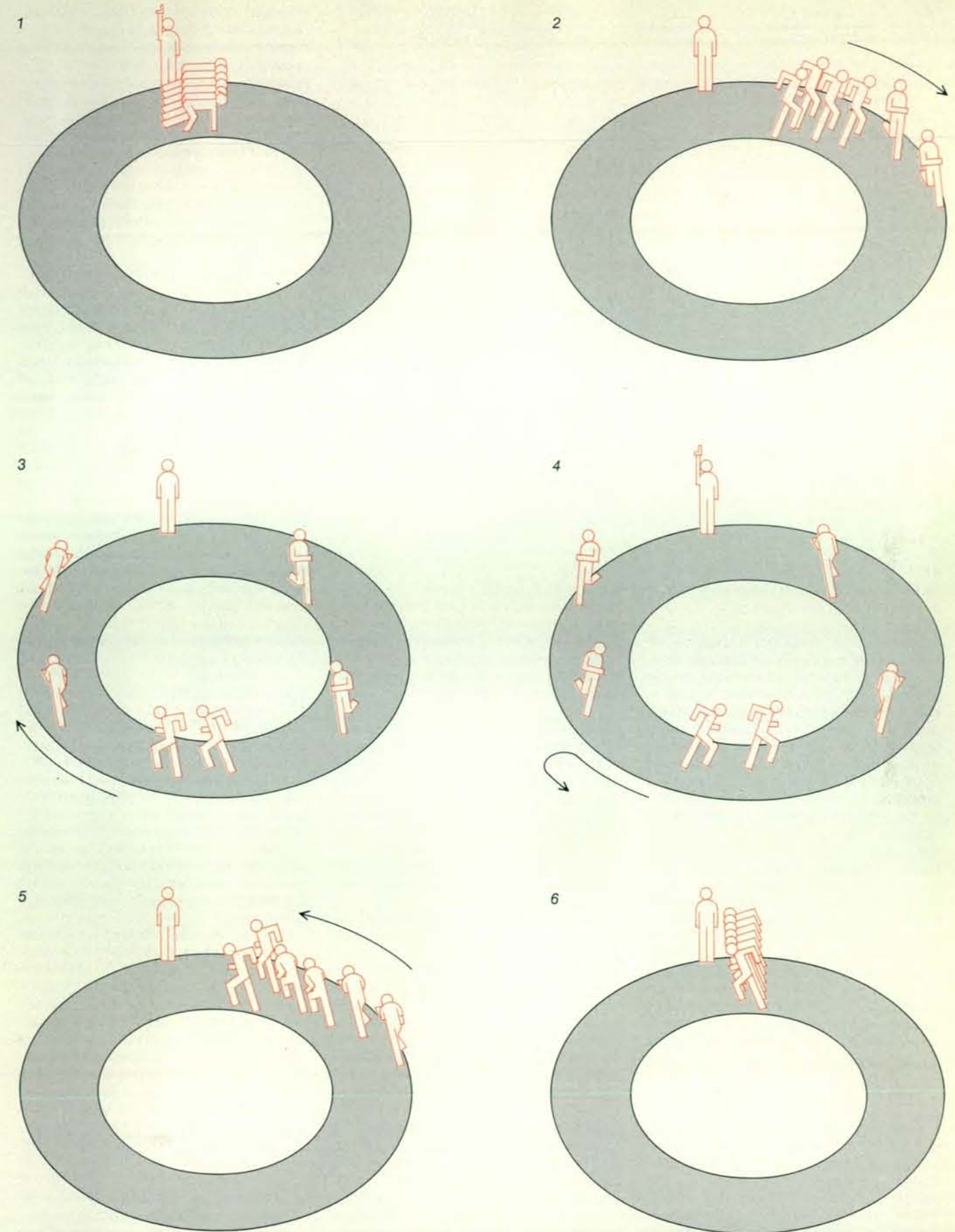
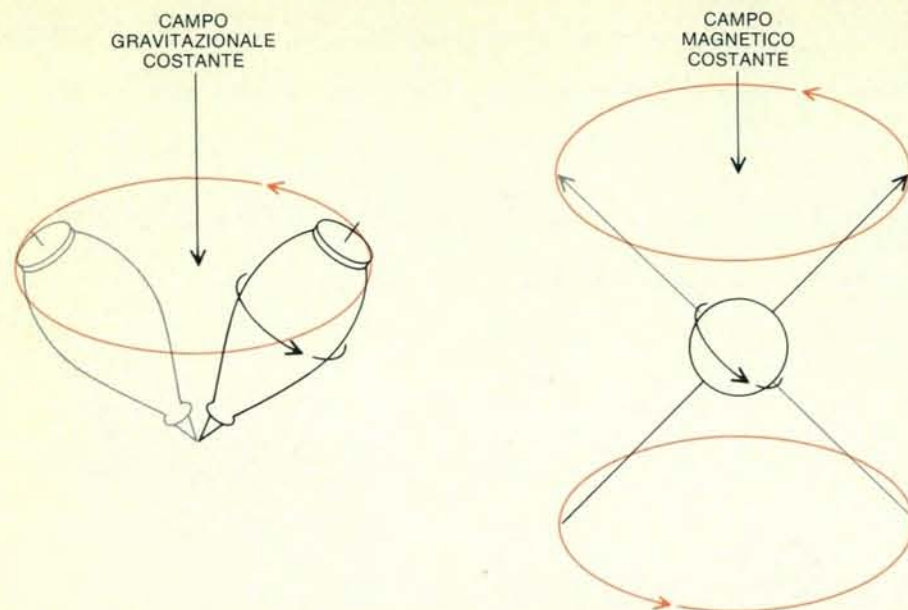
Se si mantiene il campo magnetico costante, gli spin dei protoni precederanno all'unisono in questo nuovo piano, somigliando in un certo senso a un gigantesco magnete in rotazione. Come un magnete rotante del genere, i protoni emettono un impulso elettromagnetico oscillante, il segnale di decadimento per induzione libera, così chiamato perché la precessione libera sincronizzata degli spin induce un segnale elettromagnetico che decade e che corrisponde alla partenza della corsa; gli assi di spin si trovano ora in uno stato di ordine dinamico.

Nel tempo questo ordine decade. Uno dei motivi del decadimento del segnale di induzione libera può ricercarsi nel fatto che il campo magnetico statico non ha esattamente la stessa intensità attraverso tutto il campione. Dato che la frequenza di precessione dei protoni dipende dall'intensità del campo esterno, un protone in una regione nella quale il campo magnetico è più intenso precederà più velocemente degli altri, allo stesso modo in cui alcuni corridori sono più veloci di altri. Gli assi di spin finiscono con l'orientarsi perciò in direzioni differenti. Essi si disperdono gradualmente, come i corridori sulla pista; gli assi di spin di quei protoni che precedono velocemente puntano in avanti rispetto agli assi dei protoni che precedono più lentamente *(si veda l'illustrazione a pagina 19)*. L'angolo tra due qualsiasi assi di spin orientati in direzioni differenti è l'angolo di fase; la grandezza dell'angolo di fase è una misura di quanto i due assi di spin siano fuori sincronizzazione. Sfasandosi, i protoni non emettono più il campo elettromagnetico oscillante, cioè il segnale di decadimento per induzione libera. Il campione si trova ora in uno stato di caos apparente.

Dopo il decadimento per induzione libera dei protoni il campione viene eccitato con il secondo impulso rf, che corrisponde al secondo colpo di pistola dello starter. Questo impulso ha la stessa frequenza del primo, ma normalmente ha durata doppia; di conseguenza il piano in cui giacciono gli assi di spin dei protoni viene ruotato di 180 gradi, finendo ancora una volta con un'orientazione perpendicolare al campo costante. È come se il piano in cui giacciono gli assi di spin fosse stato capovolto o riflesso in uno specchio.

In seguito al primo impulso rf, gli sfasamenti tra gli spin che precedono più velocemente e quelli che precedono più lentamente sono progressivamente diventati maggiori. Subito dopo il secondo impulso, che fa rovesciare il piano in cui giacciono gli assi, gli angoli di fase tra i vari assi sono gli stessi di prima, ma le posizioni relative degli assi che precedono più velocemente e di quelli che precedono più lentamente sono invertite. In altre parole, prima del secondo impulso gli assi che precedono più velocemente si sono gradatamente portati più avanti di quelli che precedono più lentamente; dopo il secondo impulso il piano in cui giacciono gli assi è stato rovesciato (o sottoposto a una riflessione speculare) e quelli che precedono più lentamente anticipano gli altri.

Gli spin dei protoni con precessione più rapida si trovano ora dietro ai più lenti, come i corridori più veloci erano dietro ai più lenti dopo il secondo segnale dello starter. Come nella corsa, gli assi di spin che precedono più velocemente finiranno per raggiungere i più lenti e si riallineeranno. In questo momento gli atomi emetteranno un altro impulso di radiazione, l'impulso di eco, il quale mostra che l'ordine perso è stato ripristinato.

Nell'effetto di eco di spin si dice che gli impulsi rf applicati sono in risonanza con gli spin dei protoni perché la frequenza degli impulsi coincide esattamente con la frequenza naturale di precessione dei protoni. Questa proprietà di inversione degli spin della radiazione risonante è la pietra miliare delle tecniche di risonanza magnetica nucleare (NMR) scoperta nel 1946 indipendentemente da Edward M. Purcell della Harvard University e dal compianto Felix Bloch, allora alla Stanford University. Nella spettroscopia NMR il ricercatore eccita un campione per determinare quali frequenze della radiazione indurranno l'inversione di spin; ogni frequenza di risonanza corrisponde a un solo spin nucleare in un particolare ambiente nucleare. Per esempio, l'intensità del campo magnetico locale può variare nelle diverse parti di una molecola perché una nube di elettroni scherma parzialmente il proprio nucleo da un campo esterno. Allorché un tecnico NMR conosce le frequenze di inversione degli spin, egli può determinare la composizione chimica del campione. Gli eco di spin sono tra le tecniche più utili per ottenere immagini NMR. I campi applicati dall'esterno si possono controllare per determinare la presenza di una frequenza di precessione su

I fenomeni di eco elettromagnetica trovano un'analogia nelle disposizioni dei corridori su una pista. Alla partenza *(1)* questi sono in uno stato altamente ordinato, sulla stessa linea. Quando lo starter spara il colpo di pistola e ha inizio la corsa, si sparpagliano *(2)* fino a quando la relazione reciproca appare disordinata *(3)*. (La relazione appare ancor più disordinata se alcuni corridori ne hanno superato altri.) Lo starter spara un altro colpo *(4)* e i corridori invertono il senso di marcia: chi era in testa ora si trova in coda. In *(6)* l'ordine originario è ripristinato.

La precessione giroscopica in un campo gravitazionale spiega la precessione di un protone in un campo magnetico. L'asse di un giroscopio inclinato rotante *(a sinistra)* descrive una circonferenza orizzontale; questo tipo di moto viene chiamato precessione attorno a una forza costante (in questo caso si tratta della forza di gravità). Allo stesso modo un protone *(a destra)*, una particella con carica elettrica e dotata di spin intrinseco, precede attorno a un campo magnetico costante.

un campione di grandi dimensioni, come per esempio un corpo umano.

C'è un altro modo per impiegare gli effetti di eco di spin nello studio delle proprietà di varie sostanze. Nella nostra analogia della pista, i corridori non arrivano tutti nello stesso istante se qualcuno di essi si è stancato e ha rallentato durante la corsa; in un certo senso qualunque variazione di velocità introduce un disordine all'interno dell'ordine nascosto. Un disordine dello stesso genere in un campione atomico potrebbe riguardare le collisioni tra gli atomi, le interazioni magnetiche tra atomi attigui o lo spostamento di un atomo da una regione dove il campo magnetico esterno è alto a una dove è più basso, con conseguente modificazione della sua velocità di precessione. Se si allunga il ritardo tra i due impulsi rf, aumenterà il disordine casuale introdotto tra gli impulsi e il segnale di eco diventerà più debole. Un fisico, o un chimico, può impiegare quindi l'intensità dell'eco, o il suo tempo di decadimento, come misura dei processi casuali che si svolgono nella materia, quali l'agitazione termica, il moto interno e la fluttuazione di campi locali.

Con l'avvento della luce laser coerente, nel 1964 il concetto di eco è stato esteso da Norman A. Kurnit, Isaac D. Abella e Sven R. Hartmann della Columbia University alle frequenze ottiche. I principi fisici fondamentali degli echi di spin e di fotone sono gli stessi: sono entrambi esempi di ordine nascosto prodotto o rivelato da radiazione coerente. Mentre, però, l'eco di spin riguarda i nuclei atomici, l'eco di fotone riguarda solitamente gli elettroni atomici. Come dimostrato da Richard P. Feynman, Frank L. Vernon, Jr., e Robert W. Hellwarth, allora al California Institute of Technology, si possono descrivere entrambe le situazioni con lo stesso formalismo matematico, che è una generalizzazione delle equazioni giroscopiche originarie di Bloch.

Gli esperimenti sopra descritti dimostrano che l'ordine nascosto all'interno di sistemi apparentemente disordinati si può talvolta rivelare. Si è anche dimostrato che certi fenomeni, quali le collisioni molecolari, possono introdurre elementi di disordine nell'ordine nascosto, facendo decadere l'intensità dell'eco. Si possono progettare esperimenti di eco capaci di annullare tali effetti casuali apparentemente irreversibili?

Sembra che questo suggerimento sia contrario all'intuizione secondo la quale le conseguenze di grande scala di eventi casuali quali le collisioni molecolari sono in teoria irreversibili. In questo caso l'intuizione è fuorviante, perché talvolta è possibile eliminare perfino gli effetti disordinanti di collisioni molecolari elastiche. Questo risultato si ottiene applicando un gran numero di impulsi incidenti in rapida successione temporale. Tali esperimenti a impulsi multipli furono eseguiti per la prima volta in risonanza magnetica nucleare (NMR) da Herman Y. Carr, ora alla Rutgers University, e da Purcell. Jan Schmidt dell'Università statale di Leida, Paul R. Berman dell'Università di New York e uno di noi (Brewer) hanno esteso in seguito le ricerche sugli impulsi multipli alla regione ottica. Descriveremo qui un esperimento di eco di fotone su un campione gassoso.

L'effetto dell'eco di fotone è in linea di principio molto simile all'eco di spin nucleare. Nell'eco di spin un impulso rf incidente entra in risonanza con gli assi dei protoni per allinearli in uno stato di ordine dinamico; questo ordine sembra decadere, ma viene ripristinato da un secondo impulso rf risonante, che fa invertire gli angoli di fase relativi dei protoni, facendoli riallineare e producendo un'eco. L'eco di fotone è analoga, tranne per il fatto che la radiazione incidente è fornita da un laser (cioè nella regione ottica), e risuona con oscillazioni della nube di elettroni che circonda ogni atomo di gas per produrre un impulso di eco.

In un tipico esperimento di eco di fotone gli atomi del gas sono in uno stato di moto termico caotico; essi si comportano come palle da biliardo, subendo collisioni elastiche che modificano la loro velocità, ma non il loro stato interno. Se gli atomi vengono diffusi elasticamente dopo essere stati eccitati dal primo impulso laser, subiscono lievi modifiche nella traiettoria e nella velocità. A causa dell'effetto Doppler, la frequenza di emissione (che è analoga alla velocità di precessione di un asse di protone) di ogni atomo eccitato viene modificata. L'insieme degli atomi non si trova più in uno stato di ordine nascosto. Nell'analogia della pista le cose vanno come se i corridori, scontrandosi, avessero modificato la loro velocità. Nel caso della NMR avviene lo stesso tipo di decadimento perché alcune molecole del campione di liquido diffondono casualmente verso regioni di differente intensità del campo magnetico.

Ritornando all'analogia della pista, supponiamo che lo starter spari più colpi di pistola in rapida successione, facendo ogni volta invertire il senso di marcia ai corridori. Anche se la velocità di un corridore si è un poco modificata (a causa di una collisione) tra un colpo di pistola e l'altro, egli si troverà ancora approssimativamente allineato con gli altri corridori quando tutti attraverseranno la linea di partenza, dato che nel breve intervallo di tempo tra i colpi di pistola non si sarà spostato di molto dalla sua posizione «ordinata».

Questa sequenza di inversioni multiple ha un effetto ancor più importante. Supponiamo che la velocità di un certo corridore sia leggermente aumentata per effetto di una collisione. Egli percorrerà ora nell'unità di tempo una distanza maggiore di quella che avrebbe percorso se non vi fosse stata alcuna collisione. Tuttavia, dato che egli inverte più volte la direzione di marcia, egli si sposterà dalla sua posizione ordinata della stessa quantità, prima in un senso, poi in senso opposto. Mediamente lo spostamento dalla sua posizione ordinata sarà perciò nullo, ed egli si troverà pressappoco allineato con gli altri corridori. In un certo senso è come se la collisione non fosse mai avvenuta.

Allo stesso modo, se un campione gassoso venisse eccitato con molti impulsi vicini nel tempo, gli spostamenti Doppler causati dalle collisioni elastiche sarebbero in media nulli. Un atomo che ha variato velocità avrà in realtà una frequenza di emissione diversa da quella dell'atomo medio, ma dato che ciascuno degli impulsi multipli farà invertire la fase di quell'atomo (allo stesso modo in cui ogni colpo di pistola fa cambiare senso di marcia a un corridore), la frequenza di emissione dell'atomo presenterà una alternanza di valori maggiori e minori della media. Mediamente gli atomi saranno eccitati all'unisono. Dal momento che essi rimangono sincronizzati, l'effetto delle collisioni elastiche è ridotto al minimo.

Dopo ognuno dei molti impulsi gli atomi ritorneranno allineati, come fanno i corridori tra i successivi colpi di pistola. Ogni volta che gli atomi si riallineano emettono un altro impulso di eco. La successione di più impulsi di eco provoca perciò una successione di molti eco, cioè un'eco tra ogni coppia di impulsi.

L'esperimento a impulsi multipli di Carr-Purcell è veramente un modo per evidenziare l'effetto comune di eco; il ri-

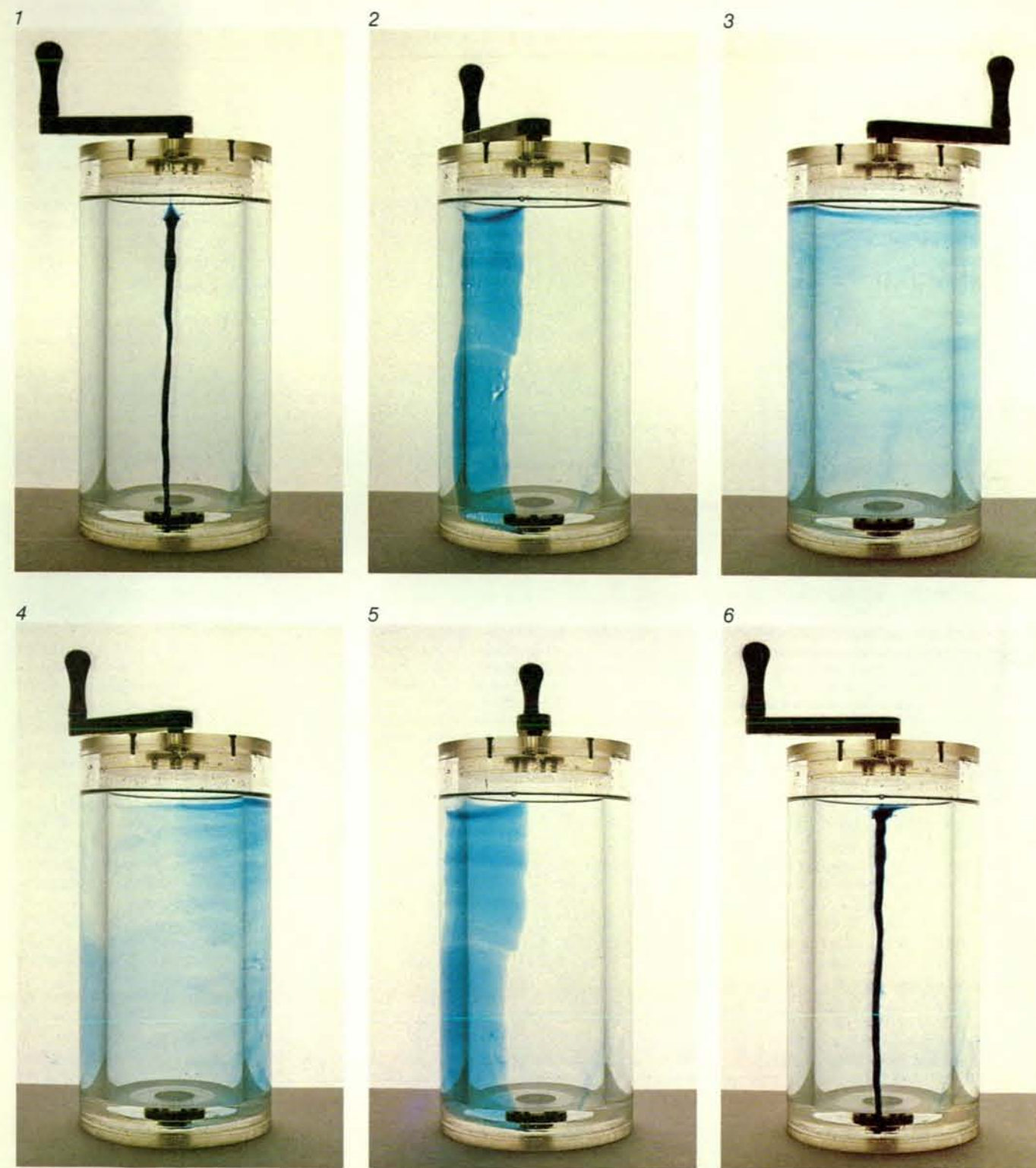
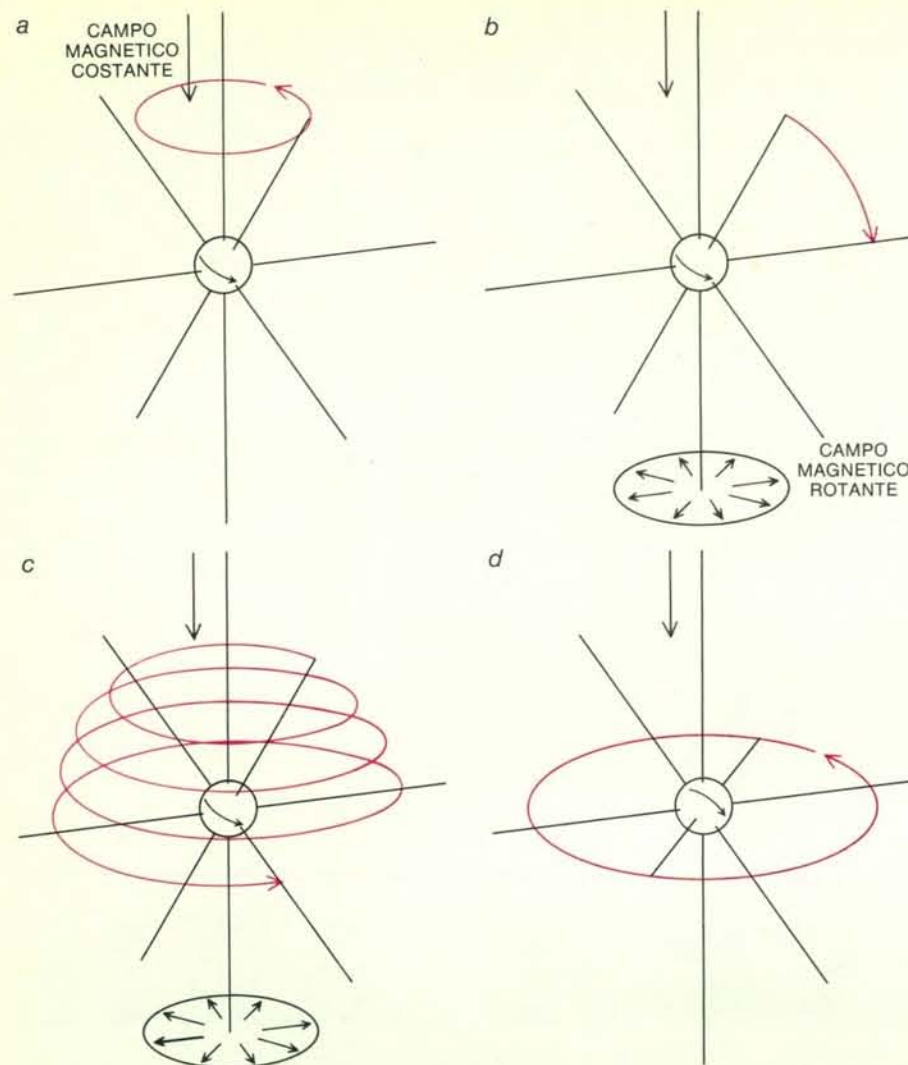
1

2

3

4

5

6

L'ordine nascosto è evidenziato da un dispositivo fatto da due cilindri trasparenti di plastica con lo stesso asse. Il volume tra i cilindri è pieno di un liquido viscoso trasparente in cui si inietta *(1)* un filetto di colorante, costituito da un allineamento inizialmente ordinato di particelle. Il cilindro esterno rimane fermo, quello interno è ruotato *(2)* finché il colorante si mescola del tutto con il liquido *(3)*. In apparenza l'allineamento iniziale scompare, ma se si inverte il senso di rotazione *(4, 5)*, le particelle si riallineano e il filetto ricompare *(6)*.

La precessione di un protone rotante soggetto sia a un campo magnetico costante sia a un campo magnetico rotante (polarizzato circolarmente) è la combinazione di due precessioni più semplici. La prima è una precessione attorno al campo magnetico costante (a, *in colore*). Se il secondo campo rotante ruota con la stessa velocità con cui precede l'asse di spin del protone, l'angolo tra i due resta costante; dal punto di vista del protone che precede, il campo rotante appare costante e l'asse precede verso il basso attorno a questo campo orizzontale (b, *in colore*). La combinazione dei due moti, vista da un osservatore esterno, è una spirale verso il basso (c, *in colore*). Il campo rotante viene «spento» quando l'asse di spin giace nel piano perpendicolare al campo costante e continua a precedere attorno al campo costante (d, *in colore*), permanendo in quel piano.

cercatore applica molti impulsi incidenti per produrre molti eco e prolungare la durata dello stato di ordine del campione.

Esiste un altro tipo di esperimento a impulsi multipli ancora più sorprendente. Chiamato «eco del sandwich magico», l'effetto è stato per la prima volta messo in evidenza da John S. Waugh, insieme a Won-Kyu Rhim e Alex Pines, a quell'epoca suoi allievi al Massachusetts Institute of Technology. Esso consiste nel sottoporre un campione a una lunga serie di impulsi trattati in modo particolare in modo da produrre una sola eco. Ciò che è peculiare dell'effetto di eco del sandwich magico è che esso può essere evidenziato in un campione che normalmente non produrrebbe alcuna eco. Senza l'applicazione di questo insolito treno di impulsi lo stato primitivo di ordine del campione non potrebbe mai essere ripristinato.

In un tipico esperimento di sandwich magico un cristallo di fluoruro di calcio viene immerso in un campo magnetico costante. Come nell'esperimento di eco di spin, al campione viene applicato un impulso rf che fa ruotare di 90 gradi gli assi di spin dei nuclei di fluoro. Il cristallo emette allora un segnale di decadimento per induzione libera come quello emesso dal campione liquido nell'esperimento di eco di spin. Una volta svanito questo segnale, si applica un altro impulso a 90 gradi, seguito immediatamente da una lunga serie di impulsi a 180 gradi in rapida successione e quindi da un altro impulso a 90 gradi. È questo il sandwich: i due impulsi a 90 gradi rappresentano le fette di pane e la serie di impulsi a 180 gradi il ripieno.

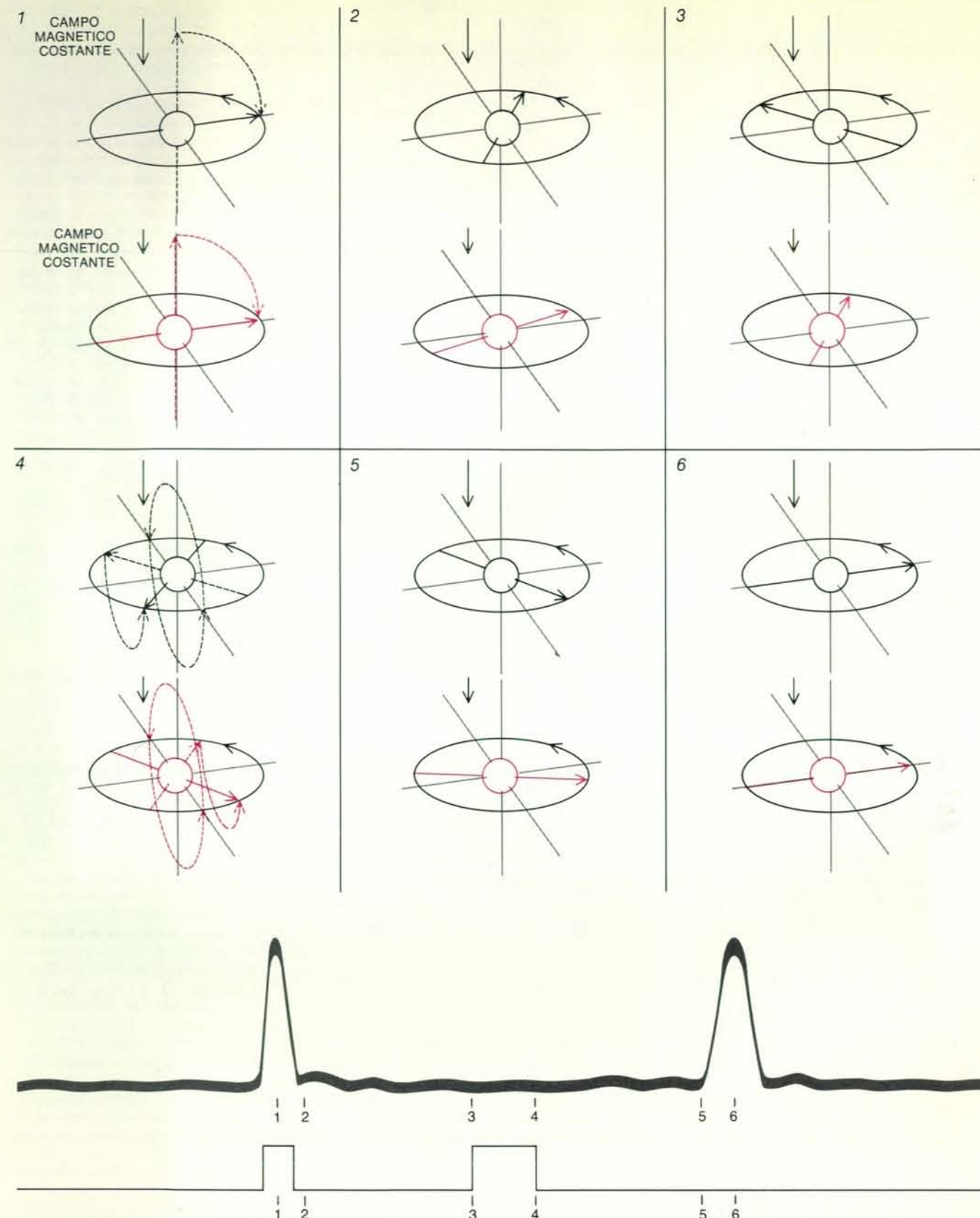
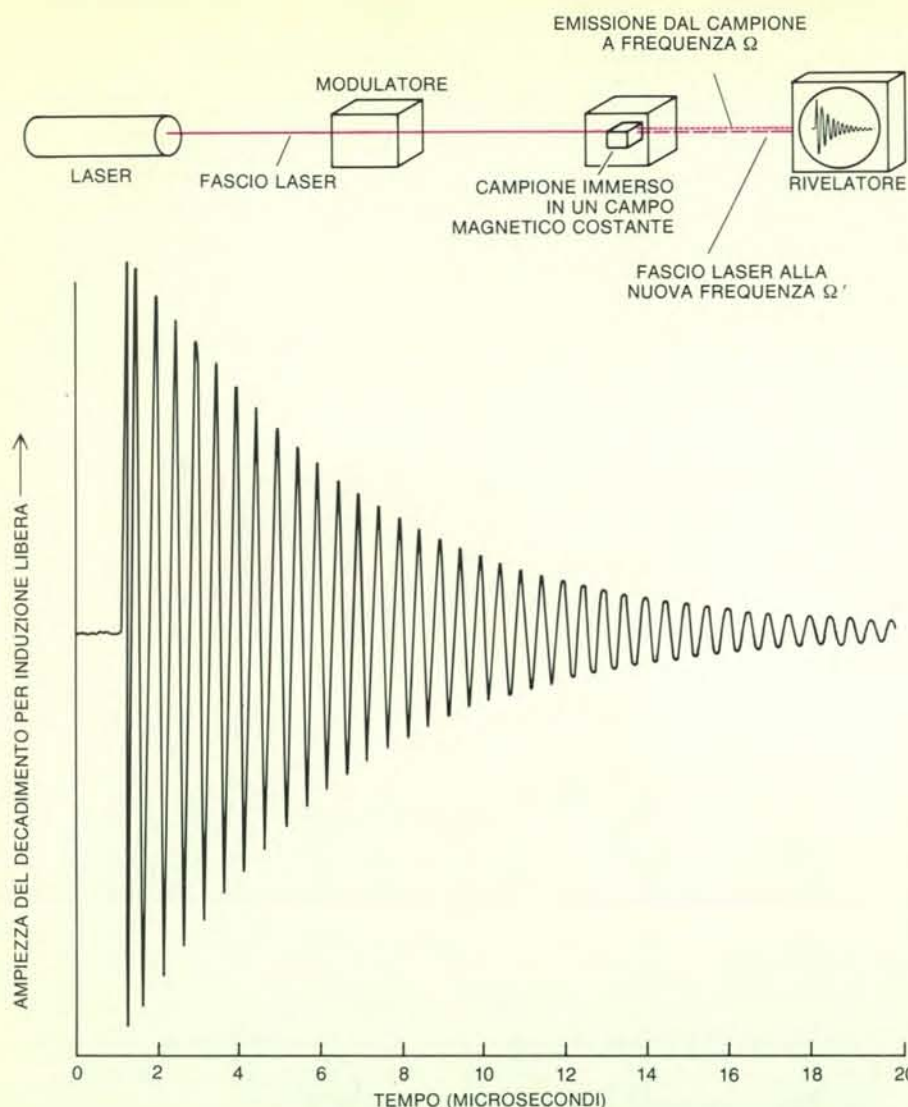
Non esiste ancora alcun modello descrittivo semplice per illustrare che cosa accade durante l'effetto di sandwich magico. Si può solo dire (secondo la descrizione matematica comunemente accettata) che in realtà il sandwich magico fa cambiare segno all'equazione di moto dei nuclei di fluoro, cioè consente proprio quella inversione della quantità di moto che Loschmidt descrisse a Boltzmann.

Perfino in un esperimento a impulsi multipli permane un certo decadimento del segnale di eco; gli echi successivi sono sempre più deboli. Questo decadimento, nel caso di un gas, è dovuto principalmente a collisioni anelastiche, collisioni sufficientemente violente per produrre variazioni quantistiche irreversibili nei livelli energetici degli atomi coinvolti. Il decadimento del segnale di eco in un esperimento a impulsi multipli è quindi una misura del tasso di collisioni anelastiche e di diffusione in un campione. Ciò significa che un esperimento a impulsi multipli può essere utilizzato per selezionare particolari tipi di interazioni atomiche da studiare senza la complicazione di processi dinamici competitivi.

Per risvegliare la memoria degli atomi v'è un altro modo ancora, nel quale non si richiede l'inversione temporale. L'idea fondamentale traspare nell'analogia della pista. È possibile che tutti i corridori procedano con la stessa velocità periferica, mantenendo il loro allineamento iniziale. Questo è il caso più semplice di tutti, ma come lo si può ottenere?

In un campione gassoso è possibile selezionare tutti gli atomi con una data velocità eccitando il campione con un fascio laser in onda continua (oc) quasi monocromatica (di una sola frequenza) con la frequenza di risonanza opportuna. A causa dell'effetto Doppler, atomi identici in moto con velocità differenti assorbiranno luce con frequenze leggermente diverse. Se la frequenza del laser è spettralmente pura, cioè se si ha realmente una sola frequenza, allora verranno selezionati e resi coerenti soltanto gli atomi con una particolare velocità. Per ritornare all'analogia della pista, è come se potessero prendere il via solo i corridori che si muovono esclusivamente a una velocità.

L'allineamento di questi atomi resi coerenti viene evidenziato in un'apparecchiatura a commutazione di frequenza laser: dopo un lungo periodo di eccitazione la frequenza del fascio laser viene portata improvvisamente a un nuovo valore non più in risonanza con il gruppo di atomi selezionati. Questa commutazione fa cessare l'eccitazione. Gli atomi selezionati in coerenza, però, si comportano ora come un gruppo di diapason identici che siano stati simultaneamente percossi: dal momento che hanno tutti la stessa frequenza di risonanza, si rinforzano l'un l'altro ed emettono all'unisono un intenso fascio coerente di luce in senso diretto. Il fascio ha tutte le proprietà della luce laser (coerenza, direzionalità e un'unica frequenza) poiché gli atomi ricordano il loro stato ordinato. Questo è l'analogo ottico del decadimento per induzione libera associato alla risonanza magnetica.

L'effetto del decadimento per induzio-

L'eco di spin è indotto da due impulsi a radiofrequenza (rf) che fanno inclinare gli assi di spin dei protoni in un campione liquido immerso in un campo magnetico costante. Sono mostrati due protoni rappresentativi; il campo magnetico nella regione del protone in nero è più intenso di quello nella regione del protone in colore. In *1* il primo impulso fa inclinare gli assi fino a portarli in un piano perpendicolare al campo costante, nel quale continuano a precedere. Poiché il campo costante differisce da una regione all'altra all'interno del campione, alcuni assi precederanno più velocemente di altri e finiranno con l'assumere orientazioni diverse (*2, 3*). Un secondo impulso rf, di durata doppia, fa ruotare il piano in cui precedono gli assi di spin fino a quando essi si troveranno nello stesso piano di prima, ma con le posizioni degli assi che precedevano più velocemente e più lentamente invertite (*4*). Gli assi con precessione più veloce si trovano ora alle spalle di quelli con precessione più lenta (*5*). Come i corridori più veloci, finiranno con il raggiungerli (*6*). Le due tracce all'oscilloscopio rappresentano la sequenza degli impulsi rf (*in alto*) e i segnali emessi dal campione quando i protoni si allineano (*in basso*), compreso l'impulso di eco (*6*).

L'apparecchiatura a commutazione di frequenza laser è stata usata da uno degli autori (Brewer) e da Ralph G. DeVoe per osservare un'altra classe di effetti di memoria atomica. Un fascio laser di frequenza Ω serve a eccitare un cristallo campione immerso in un campo magnetico costante *(al centro)*. Quindi la frequenza del laser è portata a un nuovo valore Ω' da un modulatore. Il campione risuonando sul primo fascio laser, emette radiazione coerente, il decadimento ottico per induzione libera, alla frequenza originaria Ω. L'emissione dal campione si combina con il fascio laser e produce al rivelatore *(a destra)* un segnale di interferenza, la cui durata indica al ricercatore per quanto tempo gli atomi del campione «ricordano» il fascio laser iniziale.

ne libera è stato scoperto per la prima volta nella NMR, nella regione a radiofrequenza da uno di noi (Hahn) e nella regione ottica dall'altro di noi e da Richard L. Shoemaker, ora all'Università dell'Arizona. Come l'effetto di eco, esso consente al fisico, o al chimico, di misurare, in molti tipi di materiali, proprietà normalmente difficili da osservare. Studiando il decadimento di varie frequenze di emissione in condizioni differenti, si può ottenere una migliore conoscenza delle interazioni all'interno delle molecole di un campione o tra di esse.

La commutazione di frequenza laser è stata introdotta al San Jose Research Laboratory della International Business Machine Corporation (IBM) da uno di noi (Brewer) in collaborazione con Azriel Z. Genack. Il metodo è stato utilizzato non solo per osservare il decadimento per induzione libera, ma anche un'intera classe di fenomeni che presentano memoria atomica. Il processo di accordare il laser sulla frequenza di risonanza di un campione atomico e poi disaccordarlo equivale in un certo senso ad applicare impulsi di luce laser; il portare un laser in risonanza con un campione per due brevi periodi di tempo equivale quindi in pratica ad applicare due brevi impulsi laser e produrrà lo stesso fenomeno di eco. La tecnica della commutazione di frequenza ha il vantaggio che i processi di commutazione possono essere temporizzati con maggior precisione e controllati da dispositivi elettro-ottici. Inoltre l'interferenza tra emissione del campione e luce laser (alla sua nuova frequenza) produce un intenso segnale di battimenti (eterodina) utilizzabile per distinguere l'emissione del campione da qualsiasi rumore di fondo.

Uno di noi (Brewer) con Ralph DeVoe della IBM ha recentemente applicato la tecnica di commutazione di frequenza laser allo studio delle equazioni giroscopiche fondamentali usate da Bloch nella sua prima descrizione della NMR. Secondo le equazioni di Bloch, il tempo di decadimento della memoria nucleare non dovrebbe dipendere dalla intensità dei campi applicati. Nel 1955 Alfred G. Redfield, con argomentazioni termodinamiche, dimostrò che queste equazioni devono essere modificate e osservò la risonanza magnetica nucleare di un metallo, scoprendo che un intenso campo a radiofrequenza può realmente allungare la durata del tempo di memoria (cioè ridurre il tasso di decadimento) attraverso un effetto di media temporale simile per certi aspetti all'inversione temporale.

DeVoe e Brewer hanno esteso le conclusioni di Redfield alla regione ottica. Per farlo essi hanno fatto ricorso a uno dei più stabili laser accordabili esistenti (la frequenza di emissione del laser si può regolare e, una volta fissata, è stabile fino a cinque parti su 10 bilioni). Con questo laser DeVoe e Brewer hanno eseguito un esperimento di decadimento per induzione libera. Il campione esaminato era un cristallo di trifluoruro di lantanio contenente come impurezze ioni di praseodimio. Un'interazione magnetica nucleare tra i nuclei di praseodimio e i nuclei vicini di fluoro faceva decadere la memoria.

Come un protone, anche un nucleo di fluoro si comporta come una carica rotante, creando il proprio campo magnetico. I campi dei nuclei di fluoro sono abbastanza intensi da far invertire i vicini nuclei di fluoro allo stesso modo in cui un impulso rf può invertire un protone. Quando un nucleo di fluoro si inverte, la conseguente variazione del campo magnetico locale è talvolta sufficientemente intensa da far invertire anche un altro nucleo di fluoro vicino. Tali sequenze casuali di inversione di spin sono comuni nel cristallo di trifluoruro di lantanio.

Quando un cristallo di trifluoruro di lantanio viene esposto a radiazione laser coerente con la giusta frequenza di risonanza, gli ioni di praseodimio si sincronizzano ed emettono radiazione coerente propria, un segnale di decadimento per induzione libera. I nuclei di fluoro che subiscono l'inversione casuale di spin possono desincronizzare gli ioni di praseodimio vicini, provocando lo smorzamento dell'emissione ottica coerente.

DeVoe e Brewer hanno misurato questo tempo di decadimento impiegando la tecnica di commutazione di frequenza laser descritta sopra. Essi hanno eccitato un campione di trifluoruro di lantanio con un laser a colorante accordabile e hanno commutato poi la frequenza del laser in modo che esso non potesse più risuonare con le impurità degli ioni di praseodimio. Ciò ha richiesto un laser di frequenza eccezionalmente stabile per poter eccitare un intervallo di frequenze strettissimo: la larghezza della riga del praseodimio è di soli 10 chilohertz, circa 10 milioni di volte inferiore alle prime misure ottiche nei solidi. Una volta commutata la frequenza laser, il praseodimio emetteva il segnale di induzione libera, che decadeva dopo circa 17 microsecondi.

È possibile annullare l'interazione magnetica tra il fluoro e il praseodimio aumentando l'intensità del laser usato per eccitare il cristallo. Un aumento dell'intensità del laser costringe gli ioni di praseodimio a spostarsi più rapidamente tra stati quantici più alti e più bassi mentre assorbono e riemettono fotoni. Ogni volta che lo ione praseodimio va incontro a un ciclo di assorbimento-emissione, l'interazione magnetica nucleare tra esso e i nuclei di fluoro circostanti cambia segno; in altre parole, l'interazione che si era manifestata tra uno ione praseodimio e un nucleo di fluoro agirà in direzione opposta dopo che lo ione praseodimio ha assorbito e riemesso un fotone. Così un nucleo di fluoro che aveva desincronizzato uno ione praseodimio rispetto agli altri ora invertirà il suo effetto su questo ione e lo riporterà nuovamente in sincronizzazione. Questo meccanismo è simile all'inversione dell'ordine di fase prodotta dal laser in un esperimento a impulsi multipli di Carr-Purcell. Se il tempo tra le successive inversioni dell'ordine di fase è minore del tempo tra gli eventi di desincronizzazione (in questo caso inversioni casuali del fluoro), allora le perturbazioni prodotte dai nuclei di fluoro sono compensate perché le loro interazioni con il praseodimio vengono invertite. Insieme ad Axel Schenzle dell'Università di Essen nella Germania Federale e a Masaharu Mitsunaga dell'IBM, DeVoe e Brewer hanno elaborato una teoria quantistica microscopica generale di questo fenomeno che estende per la prima volta le considerazioni termodinamiche di Redfield alla regione ottica dello spettro.

La tecnica della radiazione a radiofrequenza impulsata, i cui principi sono noti da quasi 40 anni, è uno strumento importante nella scienza e in medicina, principalmente nella formazione di immagini del corpo per mezzo della risonanza magnetica nucleare e nell'analisi strutturale di composti chimici e dello stato solido. Con lo sviluppo di laser estremamente precisi e stabili, questi metodi sono oggi stati estesi alla regione ottica.

Questi fenomeni di memoria atomica avrebbero divertito Loschmidt perché dimostrano che alcuni tipi di decadimento, anche quello causato da collisioni casuali, si possono invertire. Tuttavia, al di là del loro fascino filosofico, i fenomeni di memoria atomica possono essere molto utili. Eliminando gli effetti di decadimento di alcuni processi, essi consentono ai fisici di studiare altri processi con maggiori dettagli, fornendo una visione più chiara della struttura e delle interazioni della materia a livello atomico.

Il riconoscimento reciproco dei neuroni embrionali

Le cellule nervose in corso di sviluppo si cercano e si interconnettono con elevata specificità e negli embrioni di insetto ciò ha luogo in apparenza lungo vie la cui superficie è marcata da molecole specifiche

di Corey S. Goodman e Michael J. Bastiani

Il cervello umano si compone di centinaia di miliardi di cellule nervose, ciascuna delle quali proietta numerose fibre lunghe, che si intrecciano e si interconnettono con notevole specificità. Uno dei principali misteri della biologia continua a essere il modo in cui, durante lo sviluppo embrionale, si stabiliscono i vari collegamenti all'interno del sistema nervoso: cioè come i singoli neuroni, o cellule nervose, si trovano e si riconoscono e stabiliscono le giuste connessioni.

Le strutture che fungono da guida per la crescita delle fibre nervose embrionali vennero identificate alla fine del XIX secolo da un pioniere della neuroanatomia, lo spagnolo Santiago Ramón y Cajal. Si tratta di prolungamenti amebiformi a clava posti all'estremità delle fibre nervose; Ramón y Cajal li chiamò «coni di crescita». Lui e, più tardi, Ross G. Harrison della Yale University si avvidero che la crescita di una fibra nervosa non è affatto casuale: un particolare cono di crescita si estende sempre lungo una particolare via per trovare e riconoscere il giusto obiettivo. I due ricercatori avanzarono l'ipotesi che i coni di crescita siano dotati di qualche squisita sensibilità chimica e i loro obiettivi siano specificati chimicamente. Quest'idea venne elaborata agli inizi degli anni sessanta allorché Roger W. Sperry del California Institute of Technology (Caltech) propose l'ipotesi della chemioaffinità. Egli così concluse: «il percorso definitivo seguito da una qualsiasi fibra riflette la storia di una serie continua di decisioni, basate su affinità differenziali tra i vari filamenti di testa, che sondano l'ambiente circostante e i diversi elementi che incontrano».

Nell'ultimo decennio si sono apprese molte cose sulla struttura dei coni di crescita e sul meccanismo del loro movimento a partire da studi su neuroni dissociati in colture di tessuti, in particolare quelle realizzate da Dennis Bray del King's College di Londra, da Paul C. Letourneau dell'Università del Minnesota e da Norman K. Wessels della Stanford University. I coni di crescita emettono numerosi prolungamenti simili a peli, chiamati filopodi (i filamenti di testa di Sperry), che raggiungono molte direzioni ed esplorano l'ambiente. I filopodi sono strutture dinamiche; nello spazio di minuti si estendono, si muovono tutt'attorno e si ritraggono. Molti entrano in contatto con la superficie di altre cellule e, se vi aderiscono debolmente, si ritraggono poi all'interno del proprio cono di crescita. Ma se, invece, vi aderiscono fortemente, la loro successiva contrazione genera una tensione che guida la punta di testa del cono di crescita verso il punto di attacco. Così i coni di crescita in coltura di tessuti possono essere guidati verso particolari superfici dall'adesione differenziale mostrata dai loro filopodi.

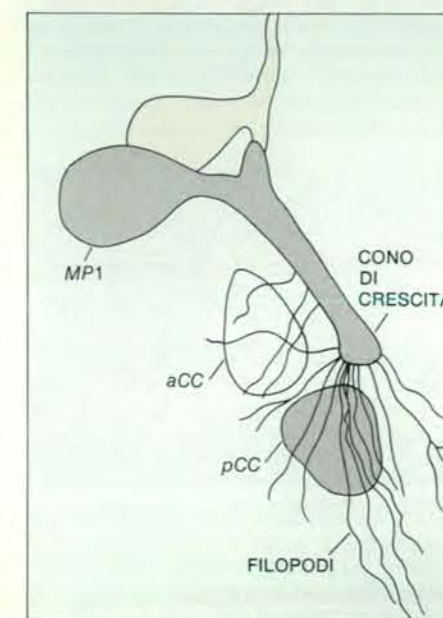
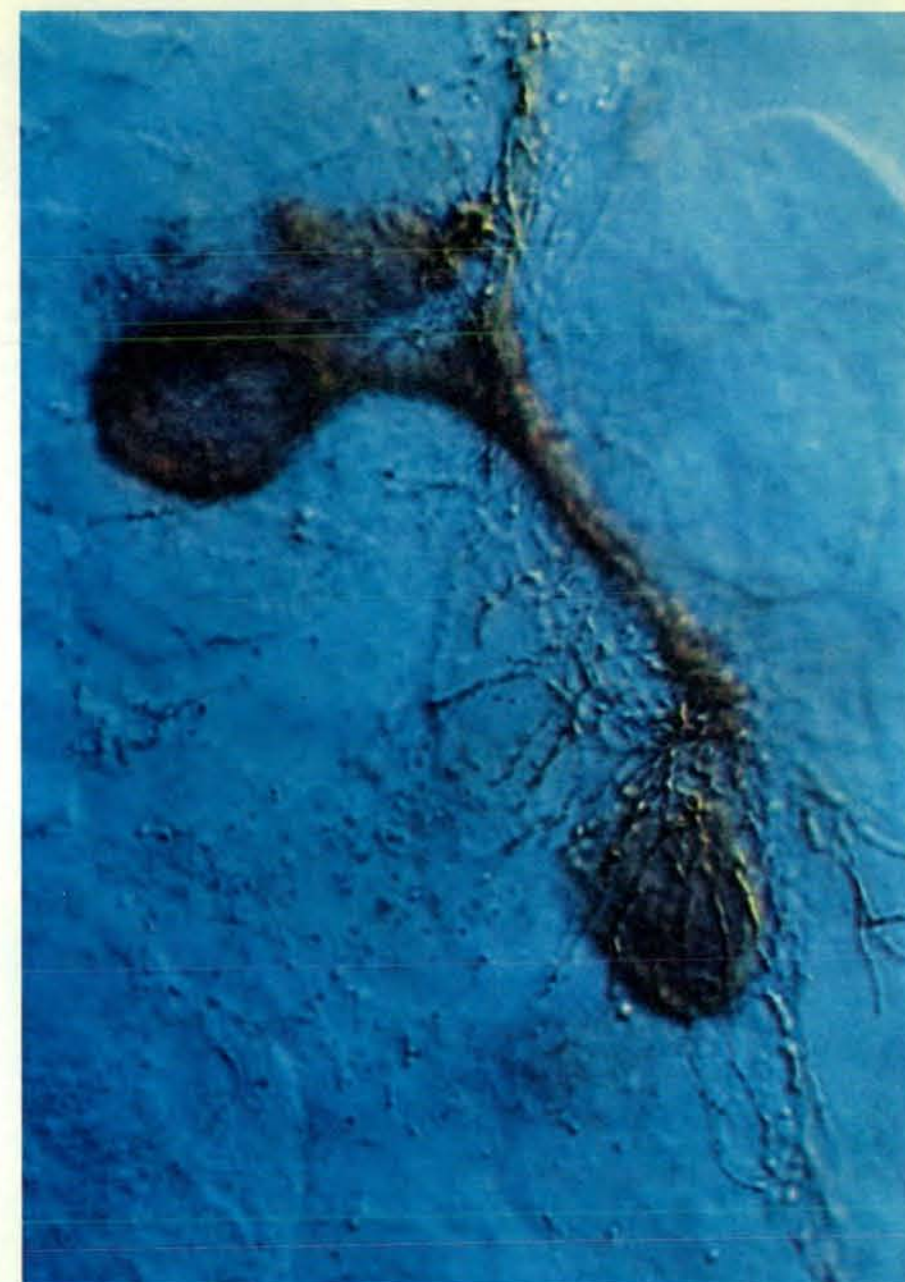
Gli interrogativi che ci siamo posti all'inizio possono ora essere inquadrati in termini più espliciti. In che modo i coni di crescita dei neuroni sono guidati verso i bersagli in un embrione in via di sviluppo? In quale misura, durante lo sviluppo, i coni di crescita e i filopodi riconoscono in modo specifico le superfici di altri neuroni, e in quale grado queste sono marcate in modo diverso? Qual è il codice molecolare dei marcatori di superficie e come viene decifrato dai coni di crescita? Per rispondere a questi interrogativi molti di noi, che sperano un giorno di riuscire a capire in che modo si sviluppino le interconnessioni nel cervello umano, hanno cominciato a studiare i cervelli di gran lunga più semplici degli invertebrati. Alla Stanford University ci siamo messi a indagare con alcuni colleghi sullo sviluppo della specificità neuronale nell'embrione di due insetti: la cavalletta *Schistocerca americana* e il moscerino della frutta *Drosophila melanogaster*.

Il sistema nervoso centrale di questi e di altri insetti include sia il cerebro (o cervello), che consta pressappoco di 50 000 neuroni, sia una catena di gangli segmentali più semplici (il ganglio è un ammasso di cellule nervose), che riflette il piano segmentato del corpo dell'animale. Ogni ganglio consta di due emisegmenti, insiemi identici di circa 1000 neuroni, uno su ogni lato. I gangli sono poi uniti da grossi fasci di assoni, i connettivi. (Gli assoni sono i processi, o prolungamenti, principali dei neuroni, quelli che inviano impulsi nervosi alle sinapsi, sedi di contatto e comunicazione tra i neuroni; i dendriti sono processi tipicamente più corti, in gran parte responsabili della ricezione di questi impulsi.) All'interno di ciascun ganglio è presente un sottile strato di corpi cellulari nei pressi della superficie ventrale (inferiore), mentre una vasta regione dorsale (superiore), il neuropilo, si trova dove tutti i prolungamenti assonici e dendritici si intrecciano e si interconnettono sinapticamente. La maggior parte dei 1000 neuroni in ogni emisegmento può essere identificata individualmente per la forma davvero unica dei loro assoni e dendriti e dello schema altrettanto unico di connessioni sinaptiche con i prolungamenti di altri neuroni. Questi «neuroni identificati» sono identici da animale ad animale all'interno di una specie.

Di tutti i sistemi nervosi degli insetti quello della cavalletta è il meglio conosciuto perché i neuroni che lo compongono sono particolarmente grossi. Un microelettrodo può facilmente essere fatto penetrare in una cellula, o per registrare impulsi nervosi e determinare la possibilità di connessioni sinaptiche, o per iniettare un colorante che penetri nel corpo cellulare e nei suoi prolungamenti, rivelando così nei particolari la struttura cellulare. Per esempio, Keir G. Pearson e collaboratori dell'Università di Alberta, in Canada, hanno studiato due neuroni, identificati nel secondo ganglio toracico (*T*2) e chiamati *C* e *G*. Questi due neuroni illustrano una parte della varietà che si riscontra nella morfologia, o struttura, della cellula e nella capacità di formare connessioni sinaptiche, rilevabile tra i 1000 neuroni che si trovano in un unico ganglio segmentale. Il neurone *C* stabilisce connessioni sinaptiche con particolari interneuroni e neuroni motori e ha parte nell'inizio del comportamento saltatorio dell'insetto. Il neurone *G*, d'altra parte, ha una funzione comportamentale molto diversa e stabilisce connessioni sinaptiche ben distinte.

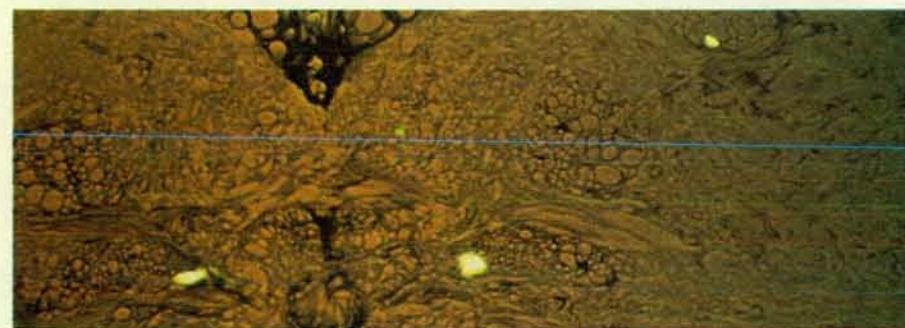
Queste differenze di funzione e di capacità di connessione si riflettono nella morfologia caratteristica di ciascun neurone, la differenza più sorprendente essendo la particolarità delle vie imboccate dall'assone e dai dendriti all'interno del neuropilo. Per esempio, in ogni cavalletta l'assone primario del neurone *G* si prolunga in avanti in un grosso fascio di fibre nervose mentre i due dendriti primari si prolungano sempre in avanti in una coppia di grossi fasci di fibre nervose disposte simmetricamente. Due principi organizzativi emergono dagli studi effettuati sui neuroni identificati nella cavalletta adulta. I processi di ciascun neurone si protendono in grossi fasci di fibre nervose secondo uno schema che è specificamente caratteristico di ogni cellula; nell'ambito dell'ambiente circoscritto da questi fasci, ogni neurone stabilisce le proprie connessioni solo con un sottoinsieme di prolungamenti neuronali che è specifico per quel neurone.

Questi principi hanno profonde implicazioni per i collegamenti che si stabiliscono nella costruzione del sistema nervoso della cavalletta durante lo sviluppo. Come nell'adulto, così anche nell'embrione il sistema nervoso consta di cellule accessibili, relativamente grandi. I gangli segmentali si sviluppano da una regione a forma di placca, il neuroepitelio, che corre dalla testa alla coda dell'embrione. All'interno di questa regione, alcune cellule si differenziano per diventare precursori dei neuroni. Nel 1976 Michael Bate, che oggi lavora all'Università di Cambridge, ha osservato che in ogni segmento del neuroepitelio uno schema stereotipato di precursori neuronali si ripete in modo praticamente identico. Un tipico segmento ha due placche laterali simmetriche, costituite da 30 neuroblasti disposti su sette file ciascuna formata da due fino a cinque cellule e un neuroblasto mediano. Sette altre cellule, i precursori della linea mediana, sono disposte appunto lungo questa linea.

Ogni precursore della linea mediana si divide una volta e dà origine a una duplice progenie di neuroni. Per contro, ogni neuroblasto funge da cellula staminale, dividendosi ripetutamente per dare luogo a una catena di cellule madri dei gangli, più piccole; ciascuna di queste ultime si divide a sua volta, generando due cellule gangliari che si differenziano e diventano neuroni. Un singolo neuroblasto fornisce così una famiglia di discendenti neuronali, da sei a circa 100, al ganglio che si sta sviluppando. Lo schema stereotipato dei circa 1000 neuroni identificati in ogni emisegmento adulto ha origine in gran parte durante lo sviluppo da uno schema parimenti stereotipato di 30 neuroblasti.

Il cono di crescita di un neurone di cavalletta designato come *MP*1 sorpassa un'altra cellula nervosa, *aCC*, e procede per andare a riconoscere la superficie di un terzo neurone che sta appunto cercando, il neurone *pCC*. Nel modo in cui questa sequenza di eventi ha potuto essere osservata, nel neurone *MP*1 è stato iniettato del colorante giallo Lucifero. Un anticorpo contro questo colorante, di colore marrone, rende visibile tale neurone. La microfotografia, interpretata dalla mappa qui a sinistra, mostra in che modo i filopodi simili a fili sottili, che si ramificano a partire dall'estremità del cono di crescita e ne guidano il corso, abbiano selettivamente preso contatto con il neurone *pCC* perché riconoscono una molecola che marca la sua superficie. Il contatto selettivo è altresì dimostrato dall'accoppiamento con il colorante: questo, iniettato nel neurone *MP*1, è passato nel neurone *pCC*, colorandolo in marrone.

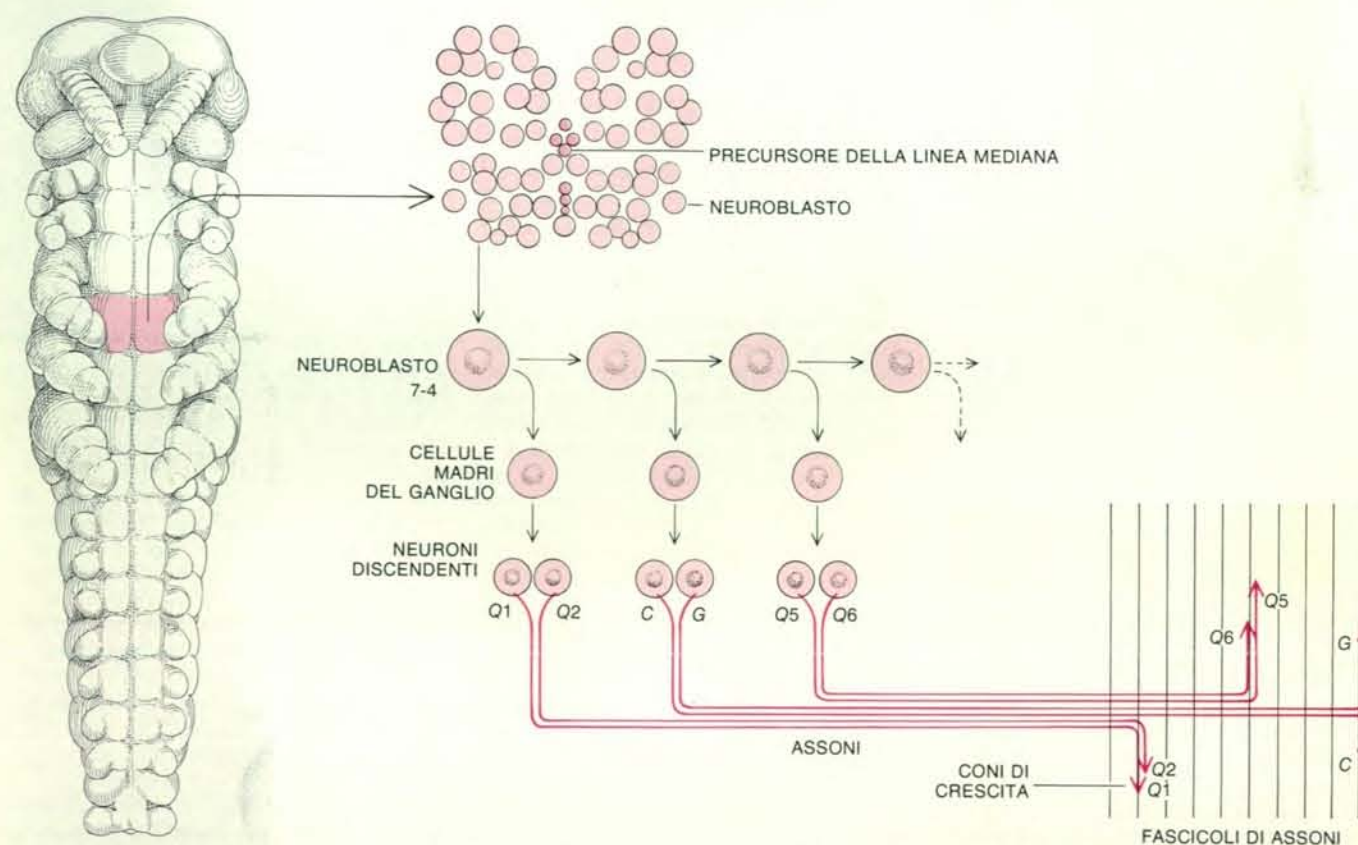
Nel 1977 Nicholas C. Spitzer e uno di noi (Goodman), mentre erano all'Università della California a San Diego, hanno cercato di stabilire se la discendenza cellulare determina alcuni aspetti del destino delle cellule. Essi si sono chiesti in primo luogo: neuroni identificati compaiono proprio in corrispondenza di specifici punti di ramificazione negli alberi genealogici dei neuroni? Per poter affrontare questa indagine, era necessario seguire i singoli neuroni dalle origini (cioè dalla divisione di particolari cellule precursori) fino allo stadio in cui ciascuno di essi non avesse esibito una forma caratteristica, che ne rivelasse la specifica identità.

Mentre i precedenti ricercatori avevano esaminato sezioni fissate, allestite mediante prelievo da embrioni, essi hanno trovato il modo di dissezionare il neuroepitelio di un embrione e di porlo sul tavolino di un microscopio composto. Il neuroepitelio è sottile e trasparente, cosicché le cellule, gli assoni e - in alcuni casi - anche i loro singoli coni di crescita hanno potuto essere evidenziati. È stato così possibile identificare singoli neuroni e seguirli dalla nascita fino alla maturazione. Facendo penetrare microelettrodi in queste cellule e nei loro prolungamenti e iniettando coloranti intracellulari, Spitzer e Goodman hanno potuto descrivere la linea di discendenza e la sequenza temporale nel differenziamento del primo discendente del neuroblasto mediano, dimostrando che un particolare neuroblasto dà effettivamente origine a una famiglia coerente di neuroni identificati.

Il neurone *G*, un «neurone identificato» (vale a dire riconoscibile in ogni cavalletta sulla base della forma e della localizzazione dei suoi prolungamenti, assoni e dendriti, e delle connessioni che stabilisce con gli altri neuroni), è reso visibile dalla fluorescenza del colorante giallo Lucifero nella microfotografia del secondo ganglio toracico del sistema nervoso centrale segmentato della cavalletta adulta, realizzata da Keir G. Pearson e John D. Steeves dell'Università di Alberta, in Canada. Il corpo cellulare, in basso a sinistra, è una delle 2000 cellule presenti nel ganglio. L'assone primario attraversa il ganglio e si prolunga in avanti *(qui verso l'alto)*; un assone secondario più piccolo si prolunga, invece, all'indietro. Due dendriti primari simmetrici si protendono in avanti verso la parte centrale del ganglio. Il neurone è ingrandito circa 100 diametri.

I prolungamenti del neurone *G* sono colorati con giallo Lucifero nella microfotografia di Keir G. Pearson, relativa a una sezione trasversale del neuropilo, la regione del ganglio toracico di cavalletta in cui assoni e dendriti di molte migliaia di neuroni si intrecciano e si connettono fra loro. L'assone primario sezionato del neurone *G* è in alto a destra, all'interno di un fascio di assoni e dendriti. I dendriti simmetrici sono visibili in basso a sinistra e al centro, in fasci diversi.

Il destino di questi neuroni potrebbe essere stabilito dalla loro ascendenza (le successive divisioni cellulari del neuroblasto e delle cellule madri del ganglio nel loro particolare albero familiare) o dalla loro posizione nel ganglio in via di sviluppo e dalle loro conseguenti interazioni con particolari cellule vicine. Si può cominciare a distinguere tra queste alternative asportando, o uccidendo, o un precursore presuntivo della cellula o le cellule vicine e notandone l'effetto. La metodica messa a punto da John E. Sulston del Medical Research Council Laboratory of Molecular Biology di Cambridge, consente l'ablazione di singoli neuroni concentrando su di essi un microfascio laser. Esperimenti simili, effettuati dai nostri colleghi Chris Q. Doe, John Y. Kuwada e Paul H. Taghert hanno dimostrato che quando un neuroblasto comincia a dividersi, la linea di discendenza svolge un ruolo di primo piano nel determinarne la progenie immediata, cioè le cellule madri del ganglio. I due neuroni fratelli prodotti dalla divisione finale di ogni cellula madre del ganglio all'inizio sono, tuttavia, equivalenti; acquisiscono destini unici solo come risultato di interazioni l'uno con l'altro prima di emettere coni di crescita. In altre parole, sia la genealogia cellulare sia le interazioni tra cellula e cellula contribuiscono a determinare il destino unico di ciascun neurone.

Avendo così delineato alcuni degli eventi precoci dello sviluppo neuronale, possiamo porre ora attenzione a quello che accade quando le cellule cominciano a emettere i loro coni di crescita. In che modo i singoli neuroni si riconoscono reciprocamente allorché i coni di crescita prendono tutta una serie di decisioni per trovare uno specifico cammino? Per tentare di rispondere a questo interrogativo con Jonathan A. Raper abbiamo studiato nei particolari il precoce sviluppo dei primi sei discendenti del neuroblasto designato come 7-4. Si tratta di neuroni *Q*1 e *Q*2, *C* e *G* (descritti prima nell'adulto) e *Q*5 e *Q*6. I coni di crescita di questi neuroni compiono, all'incirca al 40 per cento dello sviluppo, delle scelte di percorso specifiche e divergenti. (L'embrione impiega 20 giorni a svilupparsi e pertanto un giorno è equivalente al 5 per cento dello sviluppo.) Con Frances C. Thomas abbiamo esaminato queste scelte iniettando nei neuroni sostanze coloranti e ricostruendone gli assoni, i coni di crescita e i contatti stabiliti dai filopodi sulla base di

I coni di crescita all'estremità dell'assone dei neuroni *Q*5 e *G*, che ha funzione di guida *(a sinistra)*, si protendono in avanti lungo i loro cammini specifici. Le due cellule sono state iniettate con perossidasi di rafano, un enzima che è stato quindi sottoposto a una reazione chimica che colora le cellule e i loro prolungamenti in marrone. Il disegno *(a destra)* copre un campo più ampio: gli assoni hanno origine dai corpi cellulari e i filopodi si irradiano dai coni di crescita, cercando l'ambiente giusto per percorrere la strada che ogni cono di crescita riconosce.

L'embrione di cavalletta *(a sinistra)* possiede 17 gangli segmentali, con uno schema praticamente identico: 30 neuroblasti su ogni lato, un neuroblasto mediano e sette precursori dei neuroni lungo la linea mediana. Mentre un precursore si divide una sola volta e dà origine a due neuroni discendenti, un neuroblasto si divide ripetutamente e genera le cellule madri del ganglio, ciascuna delle quali si divide poi in due neuroni fratelli. Il neuroblasto designato come 7-4 ha qualcosa come 100 neuroni discendenti; l'origine dei primi sei viene mostrata qui. Il cono di crescita primario di ognuno di essi segue un percorso specifico allorché riconosce un fascicolo (fascio di assoni) e si unisce a esso *(a destra)*.

Impalcatura ortogonale formata dai fascicoli di assoni in un singolo segmento dell'embrione di cavalletta. Essa consiste essenzialmente di due schiere di fascicoli longitudinali (una per parte), connesse da fascicoli trasversali che uniscono i due lati simmetrici o emisegmenti. A questo stadio di sviluppo, circa 100 neuroni in ogni emisegmento hanno proteso i loro assoni, che si sono reciprocamente associati per formare all'incirca 25 fascicoli longitudinali. La maggior parte dei coni di crescita si trova in questo ambiente, circondata dagli assoni di altri neuroni.

microfotografie elettroniche prese in successione di sottili sezioni di tessuto embrionale.

I coni di crescita vengono a trovarsi in una selva di assoni appartenenti a neuroni già differenziati, un centinaio dei quali in ogni emisegmento ha già, a questo stadio, proteso i propri assoni nel neuropilo che si sta sviluppando. Gli assoni sono raggruppati in fascicoli, o piccoli fasci, trasversali e longitudinali. All'interno di ogni segmento questi fascicoli formano una impalcatura ortogonale pressoché quadrata, in modo che la catena dei gangli segmentali assuma una configurazione a scala di corda. Inizialmente i coni di crescita dei primi sei discendenti del neuroblasto 7-4 si comportano allo stesso modo: si estendono verso il lato opposto del neuropilo riunendosi a fascio sul fascio trasversale singolo di assoni, che costituisce lo «scalino» della scala a corda (o accrescendosi selettivamente lungo di esso). Quando, però, raggiungono l'altro lato del neuropilo, ogni cono di crescita sceglie di seguire un diverso fascicolo longitudinale.

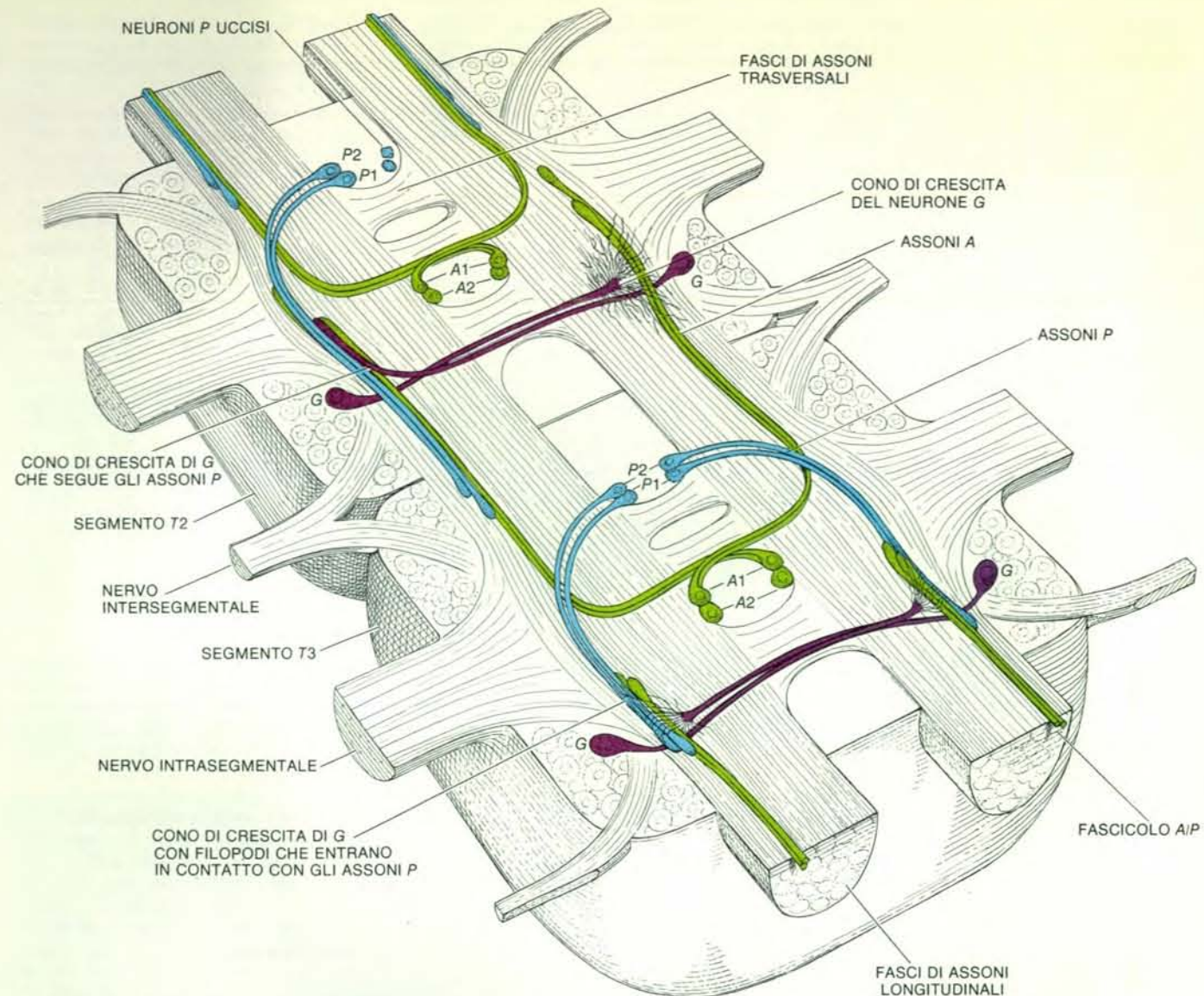
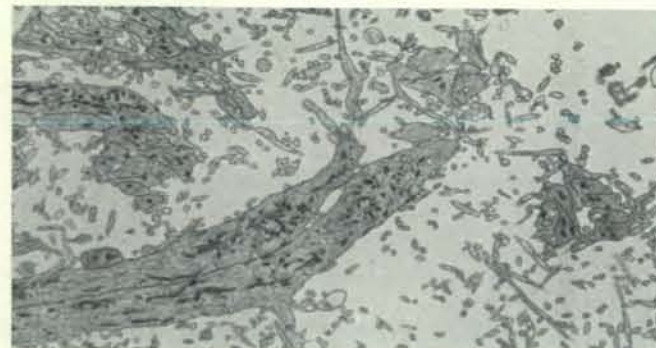
A questo punto, il cono di crescita del neurone *G* irradia abbondanti ciuffi di filopodi, che entrano in contatto con la superficie di circa 25 diversi fasci di assoni, i quali includono a questo stadio circa 100 assoni. Anche se il cono di crescita di *G* ha accesso a quasi tutte queste diverse superfici neuronali, invariabilmente sceglie di unirsi a un particolare fascio ben distinto, composto dagli assoni di quattro neuroni: *A*1, *A*2, *P*1 e *P*2. All'interno di questo fascio composto da assoni *A* e *P*, i due assoni *P* strettamente giustapposti si avvolgono attorno ai due assoni *A*. Ma ancora l'apice del cono di crescita di *G* sembra strettamente associato agli assoni *P*, non agli assoni *A*. È evidente che *G* è in grado non solo di distinguere il fascicolo degli assoni *A* e *P* dagli altri 25 fascicoli, ma anche, all'interno di quel fascicolo, di distinguere gli assoni *P* dagli assoni *A*.

Sono proprio i filopodi del cono di crescita di *G*, in particolare, che riconoscono per primi gli assoni *P*. Essi aderiscono selettivamente alla superficie dei due assoni *P* e si avvolgono attorno a essi. A questo stadio di sviluppo, si protendono lateralmente dagli assoni per tutta la loro lunghezza e così pure dai loro coni di crescita. Con nostra sorpresa abbiamo trovato che, all'interno del groviglio di filopodi che riempie buona parte del neuropilo embrionale, i filopodi di *G* possono riconoscere non solo gli assoni *P*, ma anche i loro filopodi. Di qualunque natura possano essere i segni di riconoscimento sulle superfici dei neuroni, essi sono chiaramente distribuiti sia sui filopodi sia sugli assoni e sui loro coni di crescita.

L'unione selettiva in fascicoli, mediata dai coni di crescita e dai filopodi, ha due ruoli nello stabilire le connessioni all'interno del sistema nervoso embrionale. In primo luogo i fascicoli fungono da guida per indirizzare i coni di crescita verso particolari regioni del sistema nervoso. In secondo luogo, all'interno del particolare ambito locale definito dai fascicoli, i neuroni scelgono i partner sinaptici. In una fase precoce dello sviluppo, prima che la massa all'interno del neuropilo appaia troppo fittamente aggrovigliata, i neuroni possono entrare in contatto l'uno con l'altro e riconoscersi a grandi distanze allorché i filopodi dell'uno si incontrano con i filopodi dell'altro. Questi contatti stabiliscono diversi schemi di unione selettiva in fasci. In seguito, quando si formeranno le connessioni sinaptiche, la distanza alla quale i neuroni possono entrare reciprocamente in contatto è probabilmente ridotta dalla maggior dimensione del neuropilo, dalla densità dei suoi prolungamenti e dalla mancanza di spazio per l'esplorazione da parte dei filopodi. Ne risulta che la scelta di partner potenziali da parte di un neurone è limitata a quei neuroni i cui prolungamenti siano raggiungibili nelle immediate vicinanze.

In che modo i filopodi riconoscono altri neuroni e identificano appropriate vie e obiettivi? Sulla base degli studi che abbiamo compiuto sull'unione in fasci da parte del cono di crescita di *G*, abbiamo proposto con Raper l'«ipotesi delle vie marcate». Quest'ipotesi presuppone che fascicoli vicini all'interno del neuropilo embrionale siano marcati, in maniera unica, da molecole di riconoscimento, distribuite sulle loro superfici. Al 40 per cento dello sviluppo, i filopodi di *G* si trovano alla portata di qualcosa come 25 fasci di assoni. Di conseguenza, l'ipotesi prevede che il fascicolo costituito dagli assoni *A* e dagli assoni *P*, e in particolare gli assoni *P*, abbiano un marcatore in superficie che non è presente in nessun altro dei fasci longitudinali vicini. Abbiamo controllato questo fatto eliminando selettivamente i neuroni *A*, i neuroni *P* o ambedue gli insiemi di neuroni, e notando l'effetto di ogni ablazione sul comportamento del cono di crescita di *G*.

Quando sono asportati sia i neuroni *A* sia i neuroni *P*, la maggior parte del centinaio circa di altri neuroni presenti nell'emisegmento si sviluppa normalmente. Il neurone *G* no. In assenza del fascicolo costituito da assoni *A* e *P*, il cono di crescita di *G* si comporta come se non fosse guidato, ramificandosi anormalmente e non mostrando un'affinità elevata per alcun altro fascicolo. Se si asportano solo i neuroni *A*, il cono di crescita si comporta normalmente, unendosi in fascio con gli assoni *P*. Asportando solo gli assoni *P*, invece, non mostra alcuna affinità per gli assoni *A* e non si unisce affatto.

Questi risultati suggeriscono che una pura e semplice prossimità degli assoni non può essere il fattore determinante principale, dato che (anche in assenza degli assoni *P*) gli assoni *A* rimangono al loro posto all'interno dell'impalcatura. I risultati depongono anche contro una

La riunione selettiva del cono di crescita del neurone *G* con il fascio di assoni dei neuroni *A* e *P*, costituito di due assoni *A* e due *P*, viene osservata in due segmenti adiacenti, ciascuno a uno stadio leggermente diverso dello sviluppo. Nel segmento in basso, i coni di crescita dei due neuroni *G* simmetrici (uno in ogni emisegmento) hanno attraversato il neuropilo e sono al punto in cui devono operare una «scelta», ossia selezionare il fascio longitudinale con cui riunirsi. I loro filopodi, avendo esplorato le superfici di circa 100 assoni in 25 fascicoli, hanno effettuato un contatto selettivo con gli assoni *P*. Dieci ore più tardi nel corso dello sviluppo, il cono di crescita di *G* è giunto fino al fascicolo *A/P* (costituito dagli assoni *A* e *P*) e si protende in avanti con esso *(in alto a sinistra)*. All'interno del fascicolo scelto, il cono di crescita si associa in maniera specifica con gli assoni *P* invece che con gli assoni *A*, il che suggerisce che sia in grado di riconoscere una molecola di superficie specifica per i neuroni *P*. Questo fatto è stato stabilito da Jonathan A. Raper nei suoi esperimenti di ablazione, in cui i neuroni *A* o *P* (o ambedue) venivano uccisi, cosicché i loro assoni non si sviluppavano più. Quando l'ablazione interessa soltanto i neuroni *A*, il cono di crescita di *G* si estende in avanti normalmente, raccogliendosi in fascio con i rimanenti assoni *P*. L'ablazione dei neuroni *P*, invece, fa comportare in modo anomalo il cono di crescita di *G*. Questo si ramifica a casaccio e non mostra particolare affinità per alcun altro fascicolo *(in alto a destra)*.

Il cono di crescita del neurone *G* si protende verso il fascicolo di assoni *A* e *P* nella microfotografia elettronica di sezione trasversale di neuropilo, ingrandita all'incirca 4000 diametri. Il cono di crescita è sospeso nel neuropilo mediante i suoi ampi contatti con il fascicolo per mezzo dei filopodi. Nello spazio di alcune ore l'apice avrà raggiunto il fascicolo e apparirà intimamente giustapposto agli assoni *P*.

qualsiasi determinazione da parte di qualche sottile meccanismo di scansione temporale. L'esperimento consolida invece, l'idea che la superficie degli assoni *P* presenti un marcatore qualitativamente diverso da quelli presenti su uno qualsiasi degli altri 100 assoni circostanti.

Il cono di crescita di *G* entra in contatto con qualcosa come 25 fasci di assoni e almeno uno di questi (il fascicolo costituito dagli assoni *A* e *P*) presenta due componenti differentemente marcati.

Abbiamo quindi pensato che, in questa parte del neuropilo, a questo stadio dello sviluppo, gli assoni fossero specificati da almeno 25, ma forse anche da 50, differenti segni di riconoscimento di superficie. Un così elevato grado di specificità e un così elevato numero di marcatori molecolari previsti ci hanno resi esitanti. Forse lo stadio di sviluppo pari al 40 per cento dello sviluppo completo, con qualcosa come 100 neuroni che interagiscono tra loro, era già troppo complesso per una analisi cellulare e molecolare precisa. Abbiamo così deciso di passare a uno stadio più precoce.

Abbiamo rivolto la nostra attenzione a uno stadio (pari al 30-35 per cento dello sviluppo totale), in cui i primi sette coni di crescita sono in grado di distinguere reciprocamente le proprie superfici, e attraverso interazioni specifiche, unirsi selettivamente così da formare i primi tre fasci longitudinali di assoni. Un fascio consiste soltanto dell'assone del neurone noto con la sigla *vMP2*; un altro comprende gli assoni dei neuroni *MP1*, *dMP2* e *pCC*; il terzo include gli assoni dei neuroni *U1*, *U2* e *aCC*. Con Sascha du Lac abbiamo preso in considerazione le scelte divergenti fatte dai coni di crescita dei neuroni *pCC* e *aCC*, i quali sono discendenti fratelli di una singola cellula madre del ganglio.

Le superfici di questi discendenti fratelli devono essere marcate in maniera diversa anche prima che i neuroni protendano i loro coni di crescita. Ciò è mostrato dal fatto che i filopodi del cono di crescita del neurone *MP1*, che ha uguale accesso ad ambedue i neuroni, aderisce selettivamente al corpo cellulare del neurone *pCC* e non quello di *aCC*. La differenza tra i due neuroni è ancora più chiara quando i coni di crescita cominciano a ingrossarsi. Il cono di crescita di *pCC* si estende in avanti a mano a mano che si raccoglie in fasci con gli assoni dei neuroni *MP1* e *dMP2*. Il cono di crescita del neurone *aCC* punta in avanti, ma invece di ingrossarsi rimane così per 10 o 15 ore. A questo punto, allorché gli assoni di *U1* e *U2* attraversano la regione per giungere alla portata dei suoi filopodi, sembra reagire: si gira, si muove verso i due assoni dei neuroni *U* e si congiunge con essi allorché si estendono all'indietro. Per controllare l'ipotesi delle vie marcate, du Lac ha provocato l'ablazione dei neuroni *U1* e *U2* con un microfascio laser. In assenza di questi due neuroni, il cono di crescita del neurone *aCC* è rimasto dov'era, puntando in avanti, senza mostrare alcuna affinità per la superficie di uno qualsiasi degli altri quattro assoni alla sua portata. Analogamente, l'ablazione dei neuroni *MP1* e *dMP2* ha fatto sì che il cono di crescita di *pCC* rimanesse com'era, puntando in avanti e non mostrando alcuna affinità per gli altri assoni.

Questi esempi di specificità squisita, uno allo stadio pari al 40 per cento dello sviluppo totale e l'altro allo stadio molto più semplice, pari al 30-35 per cento, ci hanno convinto che sulle superfici degli assoni embrionali sono espresse molte differenti molecole di riconoscimento e che i neuroni in fase di crescita sono guidati verso i loro obiettivi dall'adesione selettiva dei filopodi agli assoni marcati. Questa osservazione ci ha portato a prevedere che piccoli sottoinsiemi di neuroni, i cui assoni si riuniscono assieme in fascicoli, condividano uguali molecole di superficie.

Un modo per trovare queste molecole di riconoscimento postulate consiste nell'iniettare cellule embrionali in un topo e lasciare che il suo sistema immunitario generi anticorpi monoclonali contro tutte le molecole della superficie cellulare. Ogni anticorpo si lega in maniera specifica a solo una di queste molecole, il suo antigene. Analizzando migliaia di anticorpi monoclonali si può sperare di trovare un anticorpo che sia specifico per qualche molecola la cui espressione temporale e spaziale possa essere correlata con le scelte fatte da un particolare cono di crescita. È probabile che una simile molecola sia tra quelle che marcano uno specifico cammino. Abbiamo deciso di lavorare con embrioni di cavallette allo stadio pari al 40 per cento dello sviluppo completo. A questo stadio le cellule nervose non stanno ancora sintetizzando neurotrasmettitori o altre proteine prodotte dai neuroni che hanno completato il differenziamento. La loro attività principale consiste nel formare coni di crescita e nel riconoscere i percorsi da seguire: speravamo così che sarebbe stato più facile generare anticorpi contro le molecole di superficie che pensavamo interessate a quei compiti.

Kathryn J. Kotrla ha iniettato in topi neuroepitelio di embrioni al 40 per cento dello sviluppo totale, ha fuso cellule di milza di topo con cellule di mieloma maligno e ha analizzato le cellule del risultante «ibridoma» per rilevarne la capacità di sintetizzare anticorpi contro gli antigeni di superficie dei neuroni. Essa ha trovato parecchi anticorpi monoclonali che riconoscono le molecole di superficie solo su piccoli sottoinsiemi di assoni riuniti in fascicoli, proprio come aveva previsto la nostra ipotesi.

Due di questi anticorpi, chiamati *Mes*-3 e *Mes*-4, si legano alle molecole di superficie che distinguono un singolo fascio longitudinale di assoni dagli altri fascicoli, sia al 40 per cento dello sviluppo totale, quando vi sono all'incirca 25 fascicoli diversi, sia al 33 per cento, quando ve ne sono solo tre. In ambedue i casi i due anticorpi riconoscono il fascicolo costituito dagli assoni dei neuroni *MP1* e *dMP2*, quello che - come già detto - viene scelto dal cono di crescita del neurone *pCC* e non da quello del neurone *aCC*. In altre parole, un singolo fascicolo è risultato diverso, da un punto di vista antigenico, da tutti gli altri alla portata di un singolo cono di crescita.

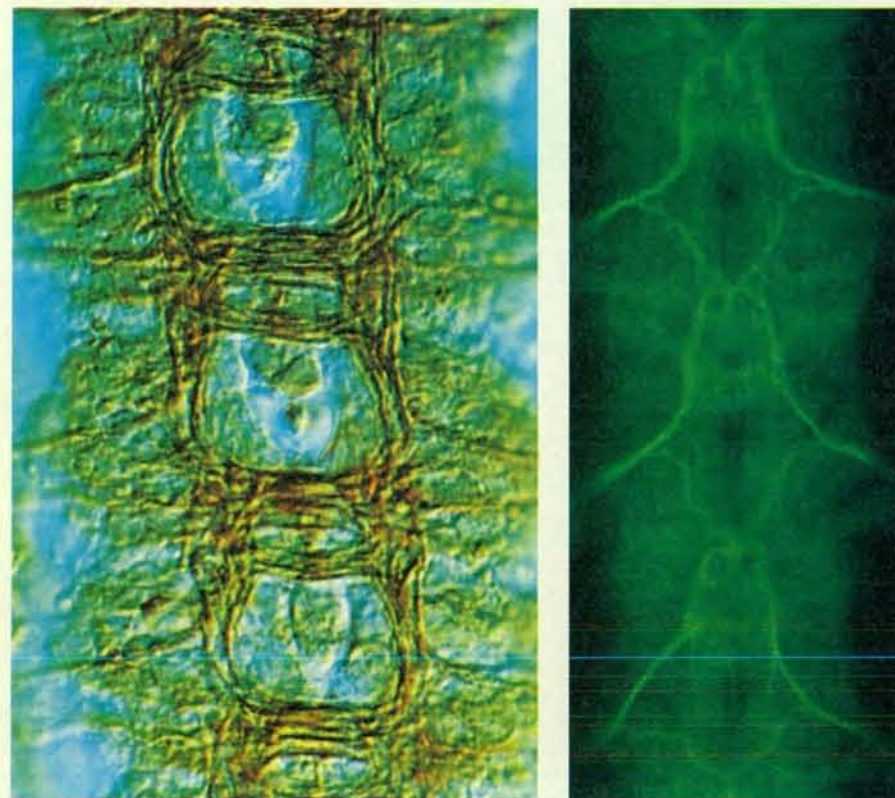
Finora abbiamo tracciato un quadro piuttosto statico di un processo di sviluppo che è, di fatto, un evento dinamico e interattivo. Per esempio, abbiamo descritto in che modo, a un certo stadio, il cono di crescita del neurone *G* e i suoi filopodi abbiano una affinità selettiva per i due assoni *P*. Tuttavia, un poco più tardi nel corso dello sviluppo, dall'assone di *G* emergono tre nuovi coni di crescita. Due di essi si estendono in avanti per diventare dei dendriti con simmetria bilaterale, mentre il terzo si protende all'indietro per diventare un assone secondario. Questi nuovi coni di crescita si uniscono in fascio con diversi altri fasci e, in ogni caso, per questi fasci il cono di crescita iniziale non mostra alcuna affinità. È chiaro che le affinità di superficie dei coni di crescita di un singolo neurone devono modificarsi con il tempo, presumibilmente come conseguenza sia delle interazioni dei coni di crescita con altre cellule sia del programma intrinseco del neurone. Per esempio, riteniamo che i cambiamenti sequenziali che si attuano nell'affinità di superficie facciano passare un cono di crescita da un fascicolo all'altro mentre si sposta verso il suo obiettivo finale. In che modo una macchina biochimica cellulare «sa» dove si trova il cono di crescita e con quali neuroni ha interagito?

La risposta può trovarsi nell'inserzione selettiva dei filopodi, un'interazione che abbiamo scoperta mentre esaminavamo delle microfotografie elettroniche per trovare la prova di un contatto tra i filopodi e superfici di cellule specifiche. Abbiamo già descritto l'adesione selettiva dei filopodi dei neuroni *G* alla superficie degli assoni e dei filopodi dei neuroni *P*. È stato interessante trovare che, invece, sono gli assoni dei neuroni *P* a entrare in contatto con il cono di crescita dei neuroni *G*. Subito dopo essi si inseriscono nel cono e, nel processo, inducono la formazione di fossette e di vescicole rivestite, strutture che normalmente si formano quando delle molecole si introducono nella cellula. L'inserimento è estremamente specifico. I filopodi di qualcosa come 40 altri coni di crescita e assoni vicini entrano in contatto con la superficie del cono di crescita di *G*, ma nessuno penetra in esso. Noi riteniamo che questo inserimento dei filopodi possa indurre un cambiamento nell'espressione delle molecole di superficie sul cono di crescita di *G*. Forse esso invia un messaggero al corpo cellulare di *G*, dove queste molecole vengono sintetizzate, segnalando in effetti: «Il tuo cono di crescita ha trovato gli assoni di *P*; preparati per la prossima scelta.»

Un esperimento con un anticorpo monoclonale dimostra questo tipo di cambiamento di specificità. L'anticorpo *Mes*-2 riconosce un antigene che si espri-

In una fase precoce dello sviluppo si modellano i primi tre fascicoli longitudinali e i coni di crescita dei neuroni fratelli *pCC* e *aCC* scelgono con quale unirsi. Il cono di crescita di *pCC* si riunisce in fascio prontamente con gli assoni di *MP1* e *dMP2* e si protende in avanti. Il cono di crescita di *aCC* sembra essere in attesa, puntando in avanti, fino a quando passano nelle vicinanze i due assoni dei gangli *U*. A questo punto si gira per unirsi a essi allorché si muovono verso la parte posteriore del corpo dell'embrione, come appare a sinistra e nella serie di disegni particolareggiati, che rappresenta sei stadi di sviluppo *(in basso)*. Sascha du Lac ha trovato che l'ablazione dei due neuroni *U* lascia arenato il cono di crescita di *aCC*, come appare a destra nel disegno grande. Il cono di crescita punta, infatti, in avanti e non si riunisce in fasci. È evidente che questi fasci precoci di assoni sono già differentemente marcati; i neuroni *pCC* e *aCC* possono già distinguerli.

L'assone del neurone *aCC* oltrepassa due altri fasci prima di riunirsi con il fascio *U*. Il montaggio della microfotografia *(a sinistra)* si riferisce a un embrione in cui ai neuroni *aCC (in alto a sinistra)* e *MP1* e a una cellula strutturale della glia *(sempre a sinistra, ma un poco più in basso)* è stata inoculata perossidasi di rafano. L'assone di *aCC* incrocia l'assone colorato in chiaro di *MP1* e quindi decorre parallelamente a esso, ma non si unisce al suo fascio, essendosi invece riunito con il fascio *U (non colorato)*. Una sottile sezione praticata in corrispondenza del sito indicato dalla freccia è ingrandita nella microfotografia elettronica *(a destra in alto)* di circa 6500 diametri.

Vie divergenti scelte dai coni di crescita dei neuroni identificati in un emisegmento. A questo stadio di sviluppo un centinaio di neuroni (qui ne sono visibili 26) hanno emesso gli assoni. Questi 100 assoni circa sono organizzati in 25 fascicoli longitudinali (qui ne sono raffigurati 10). I cerchi pieni sono tutti corpi cellulari, mentre le frecce sono i coni di crescita degli assoni primari. Lo schema interessa parti di due segmenti, dimodoché alcuni neuroni sono rappresentati più di una volta. I neuroni affini (che discendono dalla stessa cellula precursore) sono nello stesso colore. Malgrado la comune discendenza e pur dovendo affrontare lo stesso ambiente, spesso divergono nel selezionare differenti fascicoli nel neuropilo che si sta sviluppando.

La disposizione ortogonale dei fasci di assoni nella cavalletta si ripete nella *Drosophila*, ma in scala più piccola. In un embrione di *Drosophila* di 12 ore, marcato con un anticorpo contro la proteina dell'assone tubulina *(a sinistra)*, si notano tre segmenti identici. Un anticorpo monoclonale prodotto da S. L. Helfand si lega a una molecola presente soltanto su alcuni neuroni di embrione di *Drosophila*. Introdotto in un altro embrione di 12 ore e marcato indirettamente con un colorante fluorescente *(a destra)*, l'anticorpo segue l'assone di *aCC* e altri assoni che si riuniscono in fascio con esso quando dal sistema nervoso centrale si estende al di fuori nel nervo intersegmentale.

me sulle superfici di solo quattro dei 1000 neuroni presenti in ogni emisegmento (due neuroni motori e due interneuroni). Inoltre, si esprime solo transitoriamente, a un particolare stadio dell'embriogenesi. I neuroni motori sono chiamati *FETi* e *SETi*. I loro assoni si riuniscono in fasci all'esterno del sistema nervoso centrale e si dirigono verso lo stesso muscolo della gamba, che innervano. Anche se questi due neuroni motori nascono da differenti neuroblasti, essi soli nella loro famiglia hanno una via assonica finale comune ed esprimono l'antigene *Mes*-2.

Quando il cono di crescita dell'assone di *FETi* incontra per la prima volta l'assone di *SETi* al margine del sistema nervoso centrale, i due non formano un fascio comune e nessuno dei due esprime l'antigene *Mes*-2. In seguito, però, dopo aver incontrato parecchie altre cellule, il cono di crescita e l'assone si incontrano di nuovo a una distanza maggiore nella gemma dell'arto che si sta sviluppando. Questa volta essi si comportano in modo diverso. Il cono di crescita del neurone *FETi* forma un fascio con l'assone del neurone *SETi* e, insieme, i due assoni si estendono verso il muscolo che rappresenta il loro obiettivo. L'antigene *Mes*-2 comincia a essere espresso su ambedue i neuroni immediatamente prima del secondo incontro. Ancora più tardi, dopo che i due assoni hanno raggiunto e innervato il muscolo in comune, diminuisce nettamente l'espressione dell'antigene *Mes*-2, che alla fine scompare.

Eravamo molto eccitati per aver scoperto la localizzazione di questo antigene solo su quattro superfici neuronali e così pure la transitorietà della sua espressione. Queste caratteristiche riflettevano il grado di specificità previsto dall'ipotesi delle vie marcate. Nel contempo ci sentivamo frustrati perché sapevamo che sarebbe stato estremamente difficile isolare una molecola così effimera e provarne la funzione. Per farlo avremmo dovuto identificare il gene che codifica per essa e quindi studiare e manipolare l'espressione di quel gene sfruttando gli strumenti della genetica molecolare. Tutto questo sarebbe stato difficile nell'embrione della cavalletta, un insetto il cui apparato genetico non è stato molto studiato. Ci siamo rivolti allora all'embrione di *Drosophila.*

Diversamente dalla cavalletta, la *Drosophila* presenta quella che i biologi definiscono una «genetica avanzata». I suoi geni e la loro espressione hanno cominciato a essere studiati nei particolari fin dagli anni venti. Sono stati identificati migliaia di ceppi mutanti, e questo ha permesso di associare particolari geni a particolari funzioni. Gli studi di genetica nella *Drosophila,* in particolare quelli effettuati da Seymour Benzer del California Institute of Technology (Caltech) e dai suoi collaboratori e studenti, hanno offerto importanti contributi alla conoscenza dello sviluppo del sistema nervoso. Ma purtroppo le piccole dimensioni e la non accessibilità dei neuroni embrionali della *Drosophila* hanno impedito di studiare il riconoscimento cellulare e la specificità neuronale al livello di risoluzione che abbiamo raggiunto nella cavalletta.

Tuttavia, siamo stati in grado di ridurre proporzionalmente i metodi messi a punto per l'embrione della cavalletta e compiere lo stesso tipo di esperimento nella *Drosophila.* In collaborazione con John B. Thomas e con Bate a Cambridge, abbiamo trovato che il sistema nervoso centrale dell'embrione precoce di *Drosophila* è una replica in miniatura di quello della cavalletta. La somiglianza è sorprendente: in ambedue gli embrioni vi sono neuroni identificati omologhi; i loro coni di crescita si estendono nelle stesse direzioni e compiono le stesse scelte quando, muovendosi attraverso la stessa impalcatura di fasci di assoni che ricorda una scala di corda, partecipano alla formazione dei fasci. Per esempio, c'è nella *Drosophila* un neurone *G.* Il suo cono di crescita si estende attraverso il neuropilo in un fascio di assoni trasversale e poi gira in avanti lungo un fascio longitudinale. Quest'ultimo consiste di due assoni *P* e di due assoni *A* e la punta del cono di crescita del neurone *G* si associa in maniera specifica con gli assoni *P,* proprio come nella cavalletta. Questi e numerosi altri risultati fanno pensare che, nell'evoluzione degli insetti, siano stati rigorosamente conservati i primi schemi di riconoscimento cellulare.

La possibilità di lavorare con neuroni ben identificati, con i loro coni di crescita e con i tipi di riunione in fasci, messi in atto nel sistema nervoso centrale dell'embrione di *Drosophila* permette di sfruttare gli anticorpi monoclonali, le tecniche del DNA ricombinante e l'analisi genetica per svelare i meccanismi di riconoscimento cellulare e di specificità neuronale. Nel nostro laboratorio Stephen L. Helfand ha prodotto anticorpi monoclonali che riconoscono piccoli sottoinsiemi di assoni che si riuniscono in fasci. Un anticorpo, per esempio, si lega in modo specifico con una molecola espressa su alcuni assoni che costituiscono un singolo fascio: versione nella *Drosophila* del fascicolo *U* della cavalletta. Nel far questo, l'anticorpo distingue tra neuroni *aCC* e *pCC,* come ci sarebbe da attendersi sulla base della nostra analisi cellulare nella cavalletta. L'antigene che riconosce è quindi un ottimo candidato quale molecola di riconoscimento cellulare, interessata nella formazione selettiva di fasci.

Con Stephen T. Crews, Denise M. Johnson, Linda McAllister e John Thomas stiamo utilizzando questi anticorpi e altre sonde molecolari nel tentativo di isolare i geni che codificano per le molecole di riconoscimento cellulare. La comprensione dei meccanismi che sono alla base del riconoscimento cellulare da parte di 10-100 neuroni nella cavalletta e nella *Drosophila* allo stadio embrionale dovrebbe fornire un modello applicabile alle 1000 paia di neuroni che si sviluppano in ogni ganglio segmentale di un insetto, e forse anche ai miliardi di neuroni che si sviluppano nel sistema nervoso dei mammiferi.

Geni batterici nelle piante

Lo studio degli scambi di materiale genetico che possono aver luogo in natura tra batteri e piante mette oggi il ricercatore in condizione di modificare intenzionalmente il patrimonio genetico degli organismi vegetali

di Paolo Costantino e Giorgio Tecce

Il miglioramento genetico di piante importanti per l'agricoltura è una attività antica quanto l'agricoltura stessa. È anzi più antica se si considera che l'uomo non ha fatto altro che accelerare a proprio vantaggio quanto, da milioni di anni, avviene spontaneamente in natura, vale a dire la selezione di mutanti e di incroci tra varietà e specie diverse. Vale per tutti l'esempio del frumento: circa 6000 anni fa avvenne un incrocio spontaneo tra varietà selvatiche, che portò a una varietà di frumento, che fu la prima a essere sottoposta a coltivazione; circa 1500 anni dopo, un'ulteriore ibridazione, questa volta nei campi coltivati, produsse il frumento che viene coltivato ancora oggi.

Con la mutagenesi sperimentale e la selezione genetica di ibridi i tempi di produzione di varietà agronomiche «migliori» si sono ovviamente accorciati. Il limite delle pur sofisticate tecniche classiche di miglioramento genetico rimane comunque il fatto che gli incroci possono essere effettuati soprattutto tra specie affini e quindi con caratteristiche simili. Inoltre, basandosi su un rimescolamento massiccio del patrimonio genetico di due organismi, le tecniche di incrocio convenzionali si dimostrano poco flessibili quando l'obiettivo consiste nel trasferire da una pianta all'altra una sola o comunque un numero limitato di caratteristiche.

Da qualche anno a questa parte la biologia molecolare ha aperto la possibilità di modificare il patrimonio genetico di quasi ogni organismo, introducendovi direttamente singoli geni. L'insieme di tecniche che permette queste manipolazioni genetiche viene definito comunemente «ingegneria genetica», o tecnologia del DNA ricombinante (si veda l'articolo *La manipolazione dei geni* di Stanley N. Cohen in «Le Scienze», n. 87, novembre 1975). Anche l'ingegneria genetica però, nonostante sia il risultato di scoperte e di tecnologie tipicamente umane, ha un riscontro in fenomeni che avvengono spontaneamente in natura.

In questo articolo si vedrà come lo studio di un fenomeno di ingegneria genetica naturale, operato da un batterio a carico delle piante, abbia portato alla possibilità concreta di un miglioramento genetico delle specie agronomiche per mezzo della tecnologia del DNA ricombinante.

Uno dei requisiti di un sistema biologico a cui poter applicare le tecniche del DNA ricombinante è quello di disporre di una cellula ospite in grado di ricevere, replicare ed eventualmente esprimere il DNA (ricombinante), fornitole dall'esterno.

Le piante superiori sono organismi molto complessi, articolati in un grande numero di tessuti specializzati, costituiti a loro volta da cellule differenziate in morfologie e funzioni diverse fra loro. Già da alcuni decenni è però possibile coltivare *in vitro* cellule vegetali e farle crescere su terreni nutritivi a composizione nota e controllata. La cellula vegetale subisce, in coltura, un rapido processo di sdifferenziamento e può moltiplicarsi indefinitamente dando in genere origine ad aggregati, i «calli». Le cellule vegetali coltivate, o meglio i protoplasti che si ottengono da esse eliminando la parete cellulosica che le protegge, possono essere considerate e manipolate entro certi limiti come organismi unicellulari: cioè sono gli ospiti nei quali introdurre il DNA ricombinante. Anche le cellule animali possono essere mantenute *in vitro* e manipolate geneticamente con DNA ricombinante (si veda l'articolo *Ingegneria genetica in cellule di mammifero* di W. French Anderson ed Elaine G. Diacumakos in «Le Scienze», n. 157, settembre 1981). Al contrario di queste cellule, però, le cellule vegetali possono rigenerare l'intero organismo da cui sono state isolate, riuscendo a esprimere la virtuale totipotenza di cui sono geneticamente dotate tutte le cellule degli organismi superiori, come ricorda James F. Shepard nell'articolo *Rigenerazione delle piante di patata* in «Le Scienze», n. 167, luglio 1982.

La possibilità di ottenere una pianta intera da una singola cellula apre la prospettiva di modificare le piante superiori mediante la manipolazione genetica di cellule isolate. Per queste manipolazioni è possibile utilizzare sostanzialmente la stessa tecnologia messa a punto e impiegata con successo sui microrganismi (si veda l'articolo *Proteine utili da batteri ricombinanti* di Walter Gilbert e Lydia Villa-Komaroff in «Le Scienze», n. 142, giugno 1980) purché si abbia a disposizione un DNA vettore che permetta il trasferimento e l'espressione del DNA estraneo nelle cellule vegetali.

La risposta alla ricerca di tale vettore è venuta in questi ultimi anni dallo studio di alcuni tumori delle piante: le «galle del colletto» (*crown gall*). Fin dall'inizio del secolo si sapeva che moltissime piante - quasi tutte le dicotiledoni e alcune gimnosperme - sono sensibili a infezioni da parte di un batterio del suolo, *Agrobacterium tumefaciens*. In seguito a questa infezione, che di solito avviene nella regione tra il fusto e le radici (cioè nel colletto), si sviluppa sulla pianta un tumore la cui morfologia dipende dal particolare ceppo di *Agrobacterium* infettante. La presenza continua dei batteri non è necessaria per la crescita delle cellule tumorali: queste possono infatti essere coltivate *in vitro* in assenza di *Agrobacterium* e si moltiplicano mantenendo indefinitamente le loro caratteristiche tumorali. Queste caratteristiche consistono nella capacità di crescere in terreni di coltura privi di ormoni vegetali (necessari invece alla crescita di cellule normali) e soprattutto nella capacità di sintetizzare composti completamente sconosciuti alle cellule vegetali normali: le opine.

Le opine sono insoliti derivati di alcuni amminoacidi, che vengono utilizzati da *Agrobacterium* come sostanze nutritive. Va sottolineato che solo *Agrobacterium* è in grado di nutrirsi delle opine sintetizzate dalle cellule tumorali: ciò dà a questo microrganismo un vantaggio selettivo su tutti gli altri microrganismi della rizosfera. Ma come riesce *Agrobacterium* a «riprogrammare», sia pure parzialmente, le cellule infettate? Per parecchi anni si è cercato di identificare un principio induttore del tumore (TIP) in svariate componenti cellulari del batterio. La data d'inizio della biologia molecolare di questo sistema va fatta risalire al 1974 quando Jeff Schell e collaboratori, a Gand in Belgio, dimostrarono in tutti i ceppi virulenti di *Agrobacterium* la presenza di plasmidi di grandi dimensioni, assenti dai ceppi non infettivi dello stesso batterio. La capacità di indurre tumori poteva essere trasferita da un batterio a un altro in seguito al trasferimento di questi plasmidi, che furono chiamati plasmidi *Ti* (*Tumor-inducing*).

Nel 1976 Schell e Robert A. Schilperoort dimostrarono che anche la sintesi e il catabolismo delle opine sono codificati da geni localizzati sui plasmidi *Ti*. Infine, nel 1977, Mary-Dell Chilton fornì la chiave per l'interpretazione molecolare dell'induzione delle galle del colletto: una piccola parte dei plasmidi *Ti* viene trasferita dal batterio alla cellula infettata. Questo DNA batterico (*T*-DNA o DNA trasferito o trasportato) causa una vera e propria trasformazione delle cellule vegetali, analoga alla trasformazione delle cellule animali da parte dei virus oncogeni (si veda l'articolo *Come i virus inseriscono il proprio DNA in quello della cellula ospite* di Allan M. Campbell in «Le Scienze», n. 103, marzo 1977).

Il trasferimento di informazione genetica da un batterio a un organismo superiore, la pianta, non ha precedenti nei sistemi biologici fino a oggi studiati e per esso è stata coniata la definizione di «colonizzazione genetica». Dal 1977 parecchi laboratori in vari paesi del mondo si sono impegnati nello studio delle modalità di inserzione, integrazione ed espressione del *T*-DNA batterico nelle cellule vegetali e l'accumulo di conoscenze sul sistema di *Agrobacterium* ha subito in questi ultimi anni un'accelerazione con pochi precedenti nella storia della biologia molecolare.

Sarà utile a questo punto ripercorrere a grandi linee le tappe dell'evoluzione di questi studi attraverso alcuni degli esperimenti che furono decisivi: ciò aiuterà a comprendere come le tecniche di biologia molecolare, di genetica, di biochimica e di biologia cellulare cooperino nella moderna analisi di un fenomeno biologico.

Le prime osservazioni di plasmidi *Ti* furono effettuate con il microscopio elettronico: le grandi molecole di DNA circolare erano completamente assenti nelle preparazioni di DNA di ceppi di *Agrobacterium* non virulenti. La dimostrazione che tutti i geni responsabili per la virulenza di *Agrobacterium* e per l'induzione di tumori sono localizzati su tali plasmidi si ebbe attraverso esperimenti di coniugazione batterica. Durante la coniugazione tra due batteri, eventuali elementi genetici extracromosomici - quali ad esempio i plasmidi - possono essere trasferiti da una cellula all'altra. In questo modo è stato possibile trasferire i plasmidi *Ti* e conferire la capacità di indurre tumori a ceppi non virulenti di *Agrobacterium*.

Con le stesse tecniche genetiche fu dimostrato che anche i geni che codificano per la sintesi delle opine (da parte delle cellule trasformate) e la loro utilizzazione (da parte dei batteri) risiedono sui plasmi-

La galla del colletto è un tumore indotto da *Agrobacterium tumefaciens*. La pianta infettata, *Kalanchoe daigremontiana*, una specie tropicale utilizzata in laboratorio per mettere in evidenza differenze morfologiche nei tumori, presenta sul fusto, dove sono stati inoculati i batteri, una massa tumorale semidifferenziata in foglie e radici (teratoma). Altri ceppi di *A. tumefaciens* inducono tumori completamente indifferenziati. La crescita del tumore è dovuta all'introduzione di un frammento di DNA batterico nelle cellule infettate. Le cellule così trasformate proliferano in maniera più o meno disorganizzata. La fotografia è di Jacques Tempé, che lavora all'Università di Parigi a Orsay.

$$HN=C(H_2N)-NH-(CH_2)_3-CH(COOH)-NH-CH(CH_3)-COOH$$

OCTOPINA

$$HN=C(H_2N)-NH-(CH_2)_3-CH(COOH)-NH-CH(COOH)-(CH_2)_2-COOH$$

NOPALINA

AGROPINA

Le opine vengono sintetizzate nelle cellule vegetali trasformate da *Agrobacterium* e non nelle cellule normali. Di alcune di esse, derivate dagli amminoacidi, viene data la struttura chimica. L'octopina e la nopalina, derivate dall'arginina, sono presenti nelle galle del colletto indotte da *A. tumefaciens*. L'agropina, derivata dalla glutammina, è invece sintetizzata nelle radici avventizie (*hairy root*) indotte da *A. rhizogenes*. La capacità di produrre opine è conferita alle cellule trasformate dal DNA batterico trasferito (*T*-DNA): il tipo di opina indotto dipende dal plasmide tumorigeno presente in *Agrobacterium*. Gli agrobatteri utilizzano le opine come fonte di carbonio e di azoto; le opine favoriscono anche il trasferimento, per coniugazione, dei plasmidi tumorigeni da un batterio all'altro e, di conseguenza, rappresentano un vantaggio selettivo per gli agrobatteri che se ne nutrono e per i plasmidi che si diffondono nella rizosfera per il loro tramite.

di *Ti*. Fu anche chiarito un ulteriore significato evolutivo delle opine: esse promuovono la coniugazione e il trasferimento dei plasmidi *Ti*. Le opine costituiscono quindi, per questi ultimi, un vantaggio selettivo dato che ne favoriscono la propagazione e la diffusione. A questo punto, dato che le cellule delle galle del colletto sintetizzano opine anche in assenza del batterio e il gene che codifica per questa sintesi si trova sul plasmide *Ti*, era logico aspettarsi che almeno una parte di esso fosse trasferita nelle cellule vegetali e fosse responsabile della loro trasformazione.

La dimostrazione di questo singolare trasferimento d'informazione genetica si ebbe attraverso esperimenti di ibridazione tra il DNA del batterio e quello delle cellule di galle del colletto. Questo tipo di esperimento, di uso universale nella biologia molecolare di questi ultimi anni, permette di verificare e localizzare eventuali omologie nelle sequenze nucleotidiche di due DNA diversi. Uno dei due DNA viene generalmente digerito con un enzima di restrizione (enzima che riconosce e taglia il DNA in particolari sequenze nucleotidiche), sottoposto a elettroforesi in modo da separare i frammenti di digestione in base al loro peso molecolare, denaturato così da separare i due filamenti che costituiscono la doppia elica, e trasferito per capillarità su un foglio di nitrocellulosa, alla quale si lega in maniera piuttosto stabile. Mettendo questo foglio a contatto con un DNA diverso, anch'esso denaturato, si formeranno doppie eliche ibride dove i due filamenti di DNA condividono sequenze nucleotidiche sufficientemente lunghe.

Con questa tecnica è stata identificata la presenza di un frammento di DNA batterico (*T*-DNA) nelle cellule vegetali trasformate. I *T*-DNA responsabili di diversi tipi di tumori sono stati da allora clonati, cioè inseriti in un plasmide di *Escherichia coli*, e amplificati in modo da averne a disposizione grandi quantità (si veda l'articolo già citato *La manipolazione dei geni* di Stanley N. Cohen in «Le Scienze», n. 87, novembre 1975) e studiati in grande dettaglio. Si è così visto che i diversi *T*-DNA hanno una parte in comune, che probabilmente è responsabile del mantenimento dello stato tumorale delle cellule, e una parte che li distingue e in cui risiedono i geni per la sintesi delle diverse opine. Sono state ricavate successivamente le sequenze nucleotidiche complete di due di questi geni e, in entrambi i casi, si è osservato che il *T*-DNA, pur provenendo da un batterio, ha molte caratteristiche tipiche del DNA di cellule superiori. In particolar modo, analoghi a quelli del DNA di cellule eucarioti sono i siti di inizio della trascrizione dell'RNA messaggero da parte della RNA-polimerasi.

La presenza di questi siti tipici degli eucarioti sul DNA batterico consente all'apparato di trascrizione delle cellule vegetali di sintetizzare RNA messaggero che potrà essere tradotto in proteine dai ribosomi delle cellule stesse. RNA messaggeri specifici del *T*-DNA sono stati identificati infatti nelle cellule vegetali trasformate. Anch'essi presentano le caratteristiche tipiche degli RNA messaggeri degli organismi eucarioti, piuttosto che quelle degli RNA messaggeri degli organismi batterici.

Molte domande restano ancora senza risposta nel sistema di *Agrobacterium*: esse riguardano, ad esempio, il trasferimento del *T*-DNA e i primissimi stadi dell'infezione: è tutto il plasmide *Ti* a entrare nelle cellule vegetali mediante un meccanismo analogo a quello della coniugazione tra batteri, o solamente la parte che sarà poi inserita nel DNA dell'ospite, cioè il *T*-DNA? Altri problemi insoluti riguardano la funzione dei vari RNA trascritti dal *T*-DNA e l'interazione di quest'ultimo con i geni della pianta che controllano la crescita cellulare, il differenziamento e l'organogenesi.

Nonostante questi aspetti per il momento ancora oscuri, è evidente che il sistema di *Agrobacterium* presenta tutte le caratteristiche per essere utilizzato nella manipolazione genetica di cellule vegetali. Il *T*-DNA riesce, infatti, a entrare in queste cellule, si integra stabilmente nel genoma e viene espresso: un gene estraneo, inserito nel *T*-DNA, dovrebbe avere buone probabilità di essere a sua volta integrato nel DNA della cellula vegetale ed espresso. Questo schema si è dimostrato efficace e un certo numero di geni è stato introdotto con successo in cellule vegetali. Di recente, la Chilton ha pubblicato su questa rivista un esauriente resoconto sulle possibilità di utilizzare i plasmidi *Ti* per modificare geneticamente le piante superiori (si veda l'articolo *Un vettore che introduce nuovi geni nelle piante* in «Le Scienze», n. 180, agosto 1983).

Vorremmo qui accennare a un altro aspetto del sistema di *Agrobacterium* che, dal punto di vista dell'ingegneria genetica, riveste un'importanza primaria. Se l'obiettivo della manipolazione genetica delle cellule vegetali è produrre piante modificate geneticamente, è necessario che le cellule, oggetto delle manipolazioni, rigenerino poi piante intere. Le cellule trasformate dal *T*-DNA dei plasmidi *Ti* sono però abbastanza diverse dalle cellule normali. Come abbiamo visto, sono infatti in grado di crescere indefinitamente *in vitro* in assenza di ormoni. Il *T*-DNA provoca profonde alterazioni dei metabolismi ormonali che si riflettono nella incapacità delle cellule tumorali di rigenerare una pianta normale. Usando *T*-DNA come vettore di un gene esogeno, il gene in questione sarebbe inserito stabilmente in una cellula, ma in una cellula tumorale molto riluttante a rigenerare l'intera pianta: le modifiche genetiche rimarrebbero quindi a livello delle cellule. I tentativi di rigenerare organismi interi da cellule di galle del colletto hanno portato a piante più o meno deformi, incapaci di crescita autonoma (i germogli venivano innestati su piante normali) e generalmente sterili.

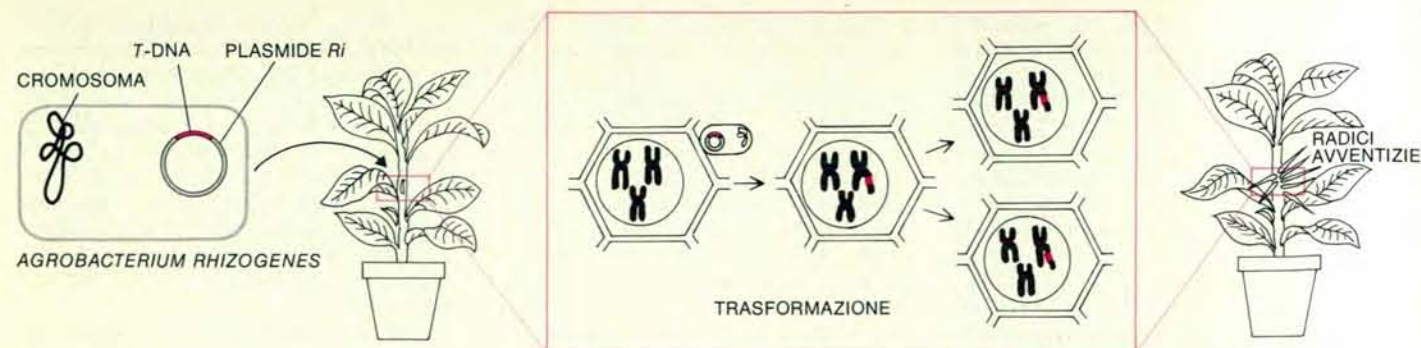
Fino a questo momento solo in poche occasioni molto particolari si è riusciti a rigenerare piante fertili e di aspetto normale da cellule trasformate da *A. tumefaciens*. In due casi si è partiti da *T*-DNA mutanti; in un terzo si è riusciti a rigenerare piante partendo da cellule ibride, ottenute fondendo cellule tumorali con cellule sane e sfruttando così la capacità rigenerativa delle cellule normali. Ci si è comunque resi conto che, per utilizzare il *T*-DNA come vettore, occorre «disarmarlo», cioè eliminare i geni responsabili della genesi tumorale e lasciare intatte le sequenze (due corti frammenti agli estremi del *T*-DNA) responsabili del suo inserimento nel DNA della cellula infettata.

Nel 1979, partendo da questa considerazione sulla difficoltà di rigenerare piante da tessuti tumorali, cominciammo a lavorare, al Dipartimento di genetica e biologia molecolare dell'Università di Roma, in collaborazione con l'ASSORENI di Monterotondo, su un organismo imparentato strettamente con *A. tumefaciens*: *Agrobacterium rhizogenes*. Questo batterio ha la particolarità di indurre la crescita di numerose radici avventizie (chiamate *hairy root*) nel sito di infezione, costituito in generale dal fusto o dalle radici di numerose specie di piante. Al momento di iniziare la nostra ricerca, nulla era noto sul meccanismo di induzione di queste radici e l'ipotesi di lavoro iniziale fu che tale meccanismo dovesse essere analogo a quello delle galle del colletto: trasformazione delle cellule da parte di un frammento di DNA batterico. Se ciò si fosse dimostrato vero, si sarebbe avuto a disposizione un ulteriore vettore per ingegneria genetica nelle piante, ma con la speranza che - dal punto di vista della rigenerazione di piante - le radici prodotte dalla trasformazione si sarebbero dimostrate più «docili» delle galle.

Il primo passo fu quindi quello di cercare di evidenziare la presenza di plasmidi di *A. rhizogenes*, ricevuti da varie parti del mondo. Per mezzo del microscopio elettronico, in collaborazione con Gioacchino Micheli del Centro acidi nucleici del Consiglio nazionale delle ricerche di Roma, riuscimmo a dimostrare che tutti i ceppi virulenti di *A. rhizogenes* contenevano grandi plasmidi, di dimensioni ancor maggiori dei plasmidi *Ti*. I plasmidi dei vari ceppi di *A. rhizogenes*, pur essendo diversi tra loro, mostravano estese sequenze nucleotidiche omologhe (evidenziate da ibridazioni molecolari) e quindi dovevano condividere molte funzioni.

Qualche tempo dopo, Frank White, a Seattle negli Stati Uniti, riuscì a dimostrare con esperimenti di coniugazione che l'infettività di *A. rhizogenes* è codificata effettivamente da questi grandi plasmidi che sono stati chiamati in seguito, in analogia con i plasmidi *Ti*, plasmidi *Ri* (*Root-inducing*). Nel frattempo, a Roma, avevamo accumulato una serie di dati che indicavano che i plasmidi *Ri* e *Ti* dovevano essere considerati in qualche modo affini, ma separati da un lungo cammino evolutivo. Di nuovo, con esperimenti di coniugazione, eravamo infatti riusciti a introdurre un plasmide *Ti* in un *A. rhizogenes*: il batterio risultante dalla coniugazione conteneva ora entrambi i tipi di plasmidi di patogeni (*Ti* e *Ri*) ed era in grado di indurre masse tumorali da cui fuoriusciva un grande numero di radici. È noto che due plasmidi molto affini non possono coesistere all'interno di uno stesso batterio: ad esempio, due plasmidi *Ti* sono incompatibili tra loro. La compatibilità dei plasmidi *Ri* e *Ti* indicava quindi una profonda divergenza evolutiva fra i due. D'altra parte ibridazioni molecolari, condotte a Roma da Gianfranco Risuleo e a Seattle da Frank White, dimostravano che una larga parte dei geni plasmidici responsabili della virulenza dei due batteri doveva essere in comune. Il *T*-DNA, cioè il DNA trasformante dei plasmidi *Ti*, non sembrava però avere nessuna omologia di sequenza con i plasmidi *Ri*. La dimostrazione che anche questi ultimi contenevano un *T*-DNA, diverso da quello responsabile delle galle del colletto, ma ugualmente trasferito nelle cellule infettate, venne nel 1982 quasi contemporaneamente dal laboratorio di Jacques Tempé, allora a Versailles, da quello di White e dal nostro. La base molecolare dell'induzione di tumori da parte di *A. tumefaciens* e di radici da parte di *A. rhizogenes* è quindi analoga; molto diverso è però il prodotto della trasformazione cellulare indotta dai due tipi di *T*-DNA.

Già da qualche anno Laura Spanò, nel

Radici avventizie vengono indotte da *Agrobacterium rhizogenes* sul fusto di una pianta di tabacco, coltivata *in vitro*. Anche in questo caso, come nelle galle del colletto, la proliferazione delle cellule è dovuta all'inserimento di *T*-DNA, originariamente localizzato su un plasmide del batterio infettante. Le cellule trasformate da *A. rhizogenes* si differenziano invariabilmente in radici.

L'infezione con *Agrobacterium rhizogenes* è prodotta in questo caso in laboratorio, mediante inoculazione di batteri in un taglio praticato sul fusto della pianta. L'agrobatterio aderisce alla parete della cellula vegetale e introduce all'interno del nucleo di questa un frammento, chiamato *T*-DNA (*in colore*), di un plasmide, il plasmide *Ri*, presente in tutti gli *Agrobacterium rhizogenes* virulenti. Le cellule risultano così «trasformate» e sono in grado di proliferare, differenziandosi in radici e mantenendo il frammento di DNA batterico in uno o più cromosomi.

nostro laboratorio, era riuscita a coltivare *in vitro*, in assenza di batteri, le radici indotte da *A. rhizogenes*. Dopo qualche incertezza iniziale ci si accorse che queste radici erano in grado di crescere indefinitamente come radici ma che, aggiungendo ormoni al terreno nutritivo, si sdifferenziavano e davano luogo a un callo indistinguibile da quelli ottenuti da radici normali. Questi calli trasformati, trasferiti su opportuni terreni nutritivi, si dimostrarono capaci di rigenerare intere piante con la stessa facilità e nelle stesse condizioni sperimentali dei calli normali.

La pianta usata nel nostro laboratorio come materiale da esperimento è il tabacco, una specie che si presta molto bene a ogni sorta di manipolazioni *in vitro*. In particolare, cellule non differenziate di tabacco, coltivate *in vitro*, sono in grado di reagire prontamente a variazioni di dosi ormonali nel terreno di coltura e rigenerano rapidamente piante intere del tutto simili a quelle ottenute da seme e cresciute in serra o in campo. Le cellule trasformate da *A. rhizogenes* (cellule «*hairy root*») si comportano da questo punto di vista come normali cellule di tabacco: le piante rigenerate nel nostro laboratorio da queste cellule trasformate sono infatti, nella maggior parte dei casi, indistinguibili da piante rigenerate da tessuti normali. La facilità nel rigenerare queste piante fu addirittura causa, all'inizio, di perplessità: i tempi e le caratteristiche di differenziamento, di organogenesi e di crescita degli esemplari così ottenuti *in vitro* coincidevano talmente con quelli delle piante rigenerate da cellule normali che pensammo di aver clonato per errore cellule non trasformate.

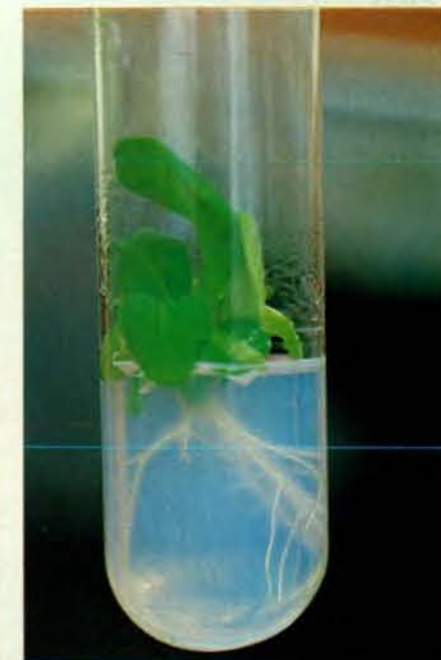
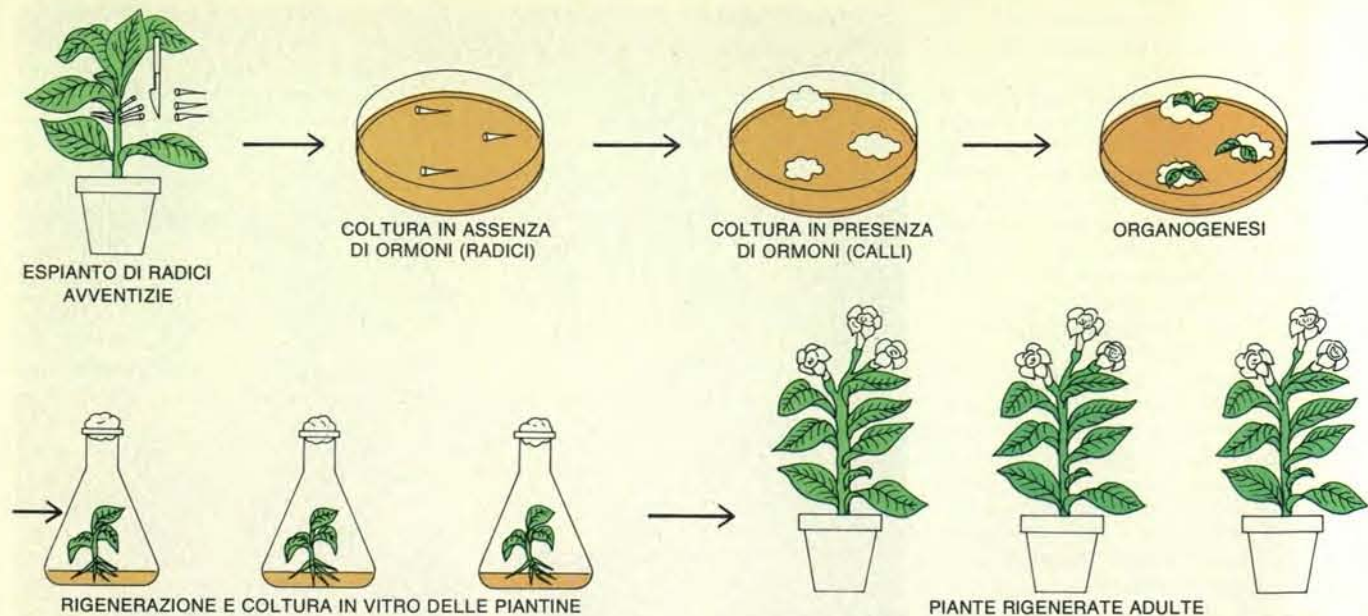
In seguito mettemmo a punto una tecnica che permette di distinguere immediatamente le cellule trasformate, contenenti il DNA batterico, da quelle normali. Questa tecnica è basata sul fatto che il *T*-DNA inserito nei cromosomi della cellula vegetale viene espresso, cioè l'informazione genetica in esso contenuta è resa operativa nella cellula ospite. Come abbiamo già accennato descrivendo le galle del colletto, uno dei prodotti di espressione del *T*-DNA nelle cellule trasformate da *A. rhizogenes* è un composto totalmente estraneo al mondo vegetale, chiamato agropina e consistente in un prodotto di condensazione tra la glutammina (un amminoacido) e il mannitolo (uno zucchero). Tali cellule sintetizzano e accumulano una grande quantità di agropina, una molecola inutile per la pianta ma che viene utilizzata dal batterio infettante come sostanza nutritiva.

L'agropina contenuta nelle cellule trasformate è facilmente evidenziabile e chiaramente distinguibile da ogni altro composto in esse presente: una elettroforesi su carta di estratti cellulari, seguita da reazione con nitrato d'argento (che colora la parte zuccherina della molecola di agropina), mette in rilievo esclusivamente l'agropina e i suoi derivati. Va sottolineato che la presenza di agropina (e delle opine in generale) è una testimonianza infallibile dell'esistenza di *T*-DNA nelle cellule e nei tessuti analizzati. Non si ha, infatti, nelle cellule vegetali normali l'informazione genetica necessaria alla sintesi di alcuna delle opine. La presenza di agropina in estratti cellulari di piante rigenerate da cellule *hairy root* indicava quindi la presenza di *T*-DNA. Ibridazioni molecolari tra il plasmide *Ri* e DNA isolato da queste piante ci confermarono in seguito che il *T*-DNA, inizialmente introdotto da *A. rhizogenes* in poche cellule all'atto dell'infezione, è mantenuto intatto attraverso le generazioni cellulari e attraverso il processo di rigenerazione. La presenza di questo *T*-DNA, al contrario di quello introdotto nelle cellule di galle del colletto da *A. tumefaciens*, non impedisce quindi alle cellule trasformate da *A. rhizogenes* di differenziarsi e di organizzarsi in una pianta funzionale. Questo fatto, come abbiamo accennato in precedenza, è molto importante ai fini dell'ingegneria genetica: con il sistema *hairy root* è possibile modificare geneticamente le cellule e rigenerare piante modificate.

Rimaneva ora da verificare un altro punto fondamentale: il *T*-DNA è presente nelle cellule sessuali delle piante trasformate da *A. rhizogenes* e, in caso affermativo, viene trasmesso alla progenie di queste piante? Va ricordato che negli organismi superiori alcune cellule, attraverso la divisione meiotica, dimezzano il proprio patrimonio genetico dando origine alle cellule sessuali (cellule uovo od oosfere e granuli pollinici). In molte piante, tra cui il tabacco, cellule uovo e granuli pollinici sono presenti nel medesimo fiore. Procedemmo quindi all'autoimpollinazione dei fiori nelle piante trasformate, mettendo poi i semi a germinare in serra. Le piante così generate, di aspetto del tutto normale, contenevano *T*-DNA e producevano agropina. Il *T*-DNA di *A.*

Nella microfotografia al microscopio elettronico è visibile (*freccia grande*) un plasmide *Ri* (*Root-inducing*) di *Agrobacterium rhizogenes*, il quale è responsabile della virulenza di questo batterio. Su di esso sono localizzati il *T*-DNA (cioè il DNA che viene introdotto nelle cellule infettate dagli agrobatteri), i geni responsabili della virulenza (che non vengono trasferiti nelle cellule infettate dagli agrobatteri) e i geni che permettono all'agrobatterio di utilizzare le opine, la cui sintesi è indotta dal *T*-DNA nelle cellule infettate. Il plasmide riprodotto è un DNA circolare extracromosomico lungo 171 000 nucleotidi. La freccia piccola indica un multimero del plasmide sintetico di *Escherichia coli* pBR322 usato come standard di lunghezza. La microfotografia è di Gioacchino Micheli del Centro acidi nucleici del Consiglio nazionale delle ricerche di Roma.

Questa coltura di radici in fase avanzata di crescita è derivata da radici avventizie indotte in una pianta di tabacco dall'infezione di *Agrobacterium rhizogenes*. L'espianto e la coltivazione *in vitro* sono realizzabili dopo avere eliminato gli agrobatteri con antibiotici. Le radici crescono molto rapidamente in assenza di ormoni vegetali. Ottenendo con i singoli apici radicali delle subcolture si possono produrre cloni da una singola cellula, che sono poi coltivabili indefinitamente.

Per aggiunta di ormoni vegetali alle colture di radici avventizie, ottenute in seguito a infezione con *Agrobacterium rhizogenes*, si ottengono colture indifferenziate (definite calli) del tipo di quella riprodotta nella fotografia a sinistra. Modificando opportunamente la composizione ormonale del terreno di coltura, si induce nel callo un'organogenesi (*al centro*), che porterà alla rigenerazione di intere piantine (*a destra*).

Lo schema riassume le fasi di espianto di radici avventizie indotte dall'infezione con *Agrobacterium rhizogenes*, la loro coltivazione in assenza e in presenza di ormoni vegetali (rispettivamente come radici e calli indifferenziati), l'organogenesi indotta con variazioni della composizione ormonale del terreno e, infine, la rigenerazione di piantine. Queste vengono fatte crescere dapprima *in vitro* e, in seguito, vengono trapiantate in vaso. Dal momento che gli apici radicali coltivati all'inizio derivano da cellule singole, anche le piante ottenute alla fine del processo di rigenerazione sono di origine clonale. Tutte le cellule di una pianta rigenerata da una delle radici trasformate da *A.rhizogenes* contengono il *T*-DNA introdotto da questo agrobatterio nella cellula infettata all'inizio. Morfologicamente questo tipo di piante (nello schema sono raffigurate piante di tabacco) non si discosta molto dalle piante rigenerate da cellule normali; si tratta di piante fertili, la cui progenie contiene ancora *T*-DNA. Utilizzando i plasmidi *Ri* per introdurre un qualsiasi gene nelle cellule vegetali, si possono ottenere piante che mantengono questo gene e lo trasmettono di generazione in generazione. È così possibile modificare geneticamente anche piante agrarie.

rhizogenes si comporta quindi come un normale gene della pianta e viene trasmesso, attraverso i complicati meccanismi di appaiamento e di segregazione di cromosomi che avvengono durante la meiosi, alla progenie sessuale delle piante in cui è inizialmente presente. Le manipolazioni genetiche effettuate in laboratorio mediante il *T*-DNA di *A. rhizogenes* dovrebbero quindi essere non solo mantenute stabilmente nelle piante rigenerate, ma anche trasmesse attraverso i semi alle generazioni successive.

Le conclusioni raggiunte nel nostro laboratorio utilizzando il tabacco sono state confermate a brevissima distanza da altri in piante diverse quali la carota, il cavolfiore, la *Nicotiana glauca* (una pianta affine al tabacco) e altre. In alcuni casi, le piante rigenerate da cellule *hairy root* si sono dimostrate un poco diverse da quelle rigenerate da cellule normali per quanto riguarda la forma delle foglie e altre caratteristiche morfologiche. Le piantine di carota trasformate hanno esternamente un aspetto normalissimo, ma sotto terra mancano qualche volta della grossa radice centrale, il fittone, per la quale vengono normalmente coltivate le carote! È anche possibile che gli effetti del *T*-DNA varino da organismo a organismo e che non in tutti i casi siano compatibili con uno sviluppo completamente normale. Ma è proprio rispetto ai problemi che riguardano l'accrescimento cellulare, il differenziamento e l'organogenesi che il sistema *hairy root* presenta caratteri del tutto unici e offre eccitanti possibilità di sviluppo.

Va ricordato che ogni singola cellula di un organismo complesso, come un animale o una pianta, contiene la stessa informazione genetica di tutte le altre, indipendentemente dalla sua funzione e dalla sua differenziazione. Molti geni però sono espressi in alcune cellule e non in altre e questo determina appunto la specializzazione morfologica e funzionale di ogni singola cellula. Che cosa determini, e come, l'espressione differenziata e modulata del patrimonio genetico sia orizzontalmente, cioè nei vari tessuti dell'organismo adulto, sia verticalmente, cioè durante lo sviluppo, costituisce uno degli argomenti più complessi e affascinanti della biologia moderna. Data l'enorme complessità del genoma delle cellule superiori è, infatti, estremamente arduo rintracciare gli elementi genetici la cui espressione, o mancata espressione, induce un particolare stato differenziativo e funzionale.

Nel sistema *hairy root* le cellule trasformate dal *T*-DNA vengono parzialmente riprogrammate da quest'ultimo. Indipendentemente dallo stato differenziativo originale della cellula (si possono infettare fusto, foglie ecc.), l'effetto del *T*-DNA è quello di indurre proliferazione e differenziamento in radici. Questo stato differenziativo è quindi indotto da qualche cosa che è codificato dal *T*-DNA. Si hanno a questo punto due possibilità: uno o più prodotti genici del *T*-DNA interferiscono con l'espressione di geni endogeni della cellula vegetale oppure uno o più prodotti genici del *T*-DNA sono direttamente responsabili della proliferazione e del differenziamento della cellula trasformata. In entrambi i casi la possibilità di studiare nei particolari la struttura e i prodotti di questo elemento genetico costituisce un'opportunità unica per entrare nel merito del differenziamento delle piante superiori.

Il *T*-DNA di un particolare ceppo di *A. rhizogenes* è stato clonato da Mirella Pomponi, nel nostro laboratorio, e la sua struttura è stata analizzata per mezzo di alcuni enzimi di restrizione. In seguito abbiamo esteso l'analisi di restrizione all'intero plasmide *Ri*, su cui è localizzato il *T*-DNA nel batterio, e ne abbiamo costruito l'intera mappa circolare analogamente a quanto era stato fatto in precedenza con i plasmidi *Ti* di *A. tumefaciens* (come riportato nel già citato articolo di M.-D. Chilton).

Attualmente stiamo «costruendo» con Maura Cardarelli, del Centro acidi nucleici del Consiglio nazionale delle ricerche, una serie di mutanti di *A. rhizogenes*, nel *T*-DNA del quale sono stati inseriti elementi genetici mobili (trasposoni) che conferiscono resistenza a svariati antibiotici. La presenza dei trasposoni in vari punti del *T*-DNA interrompe la continuità fisica dei geni, compromettendone l'espressione. La mancata espressione di un gene si rifletterà in una variazione di qualche caratteristica del sistema *hairy root*: in questo modo sarà possibile risalire alle funzioni dei vari geni del *T*-DNA che - come abbiamo visto - sono in grado di controllare la crescita e il differenziamento delle cellule. In un momento successivo (e con un po' di fortuna!) si potrà risalire ai geni che controllano le stesse funzioni nelle cellule normali.

Nel corso degli esperimenti che ci hanno portato a individuare il *T*-DNA nelle cellule trasformate da *A. rhizogenes* ci eravamo infatti accorti che, all'interno del DNA di piante normali che non sono mai state in contatto con *A. rhizogenes*, esistono sequenze abbastanza simili al *T*-DNA. Inizialmente abbiamo pensato a qualche artefatto sperimentale, ma la riproducibilità dei dati in esperimenti diversi ci ha convinto che eravamo di fronte a un risultato inaspettato, che avrebbe aperto nuove aspettative di ricerca (nulla di simile era mai stato osservato nelle galle del colletto).

Il *T*-DNA di *A. rhizogenes* è quindi strutturalmente simile al DNA che fa parte del patrimonio genetico delle piante normali. A che cosa serve questo DNA? La somiglianza con il *T*-DNA è casuale o si tratta di DNA «fisiologico», che ha le stesse funzioni di quest'ultimo nel normale sviluppo delle cellule?

Alcuni esperimenti preliminari, condotti da Maria Luisa Mauro nel nostro laboratorio, indicherebbero che questo DNA cellulare (la cui presenza è stata in seguito confermata da White in *Nicotiana glauca*) è trascritto in RNA simili a quelli prodotti dal *T*-DNA. Sono questi prodotti di trascrizione fisiologici in qualche modo coinvolti con la crescita e il differenziamento delle cellule vegetali? E ancora, qual è l'origine evolutiva del *T*-DNA che - come abbiamo visto - somiglia più a un DNA di organismo superiore che non a un DNA batterico?

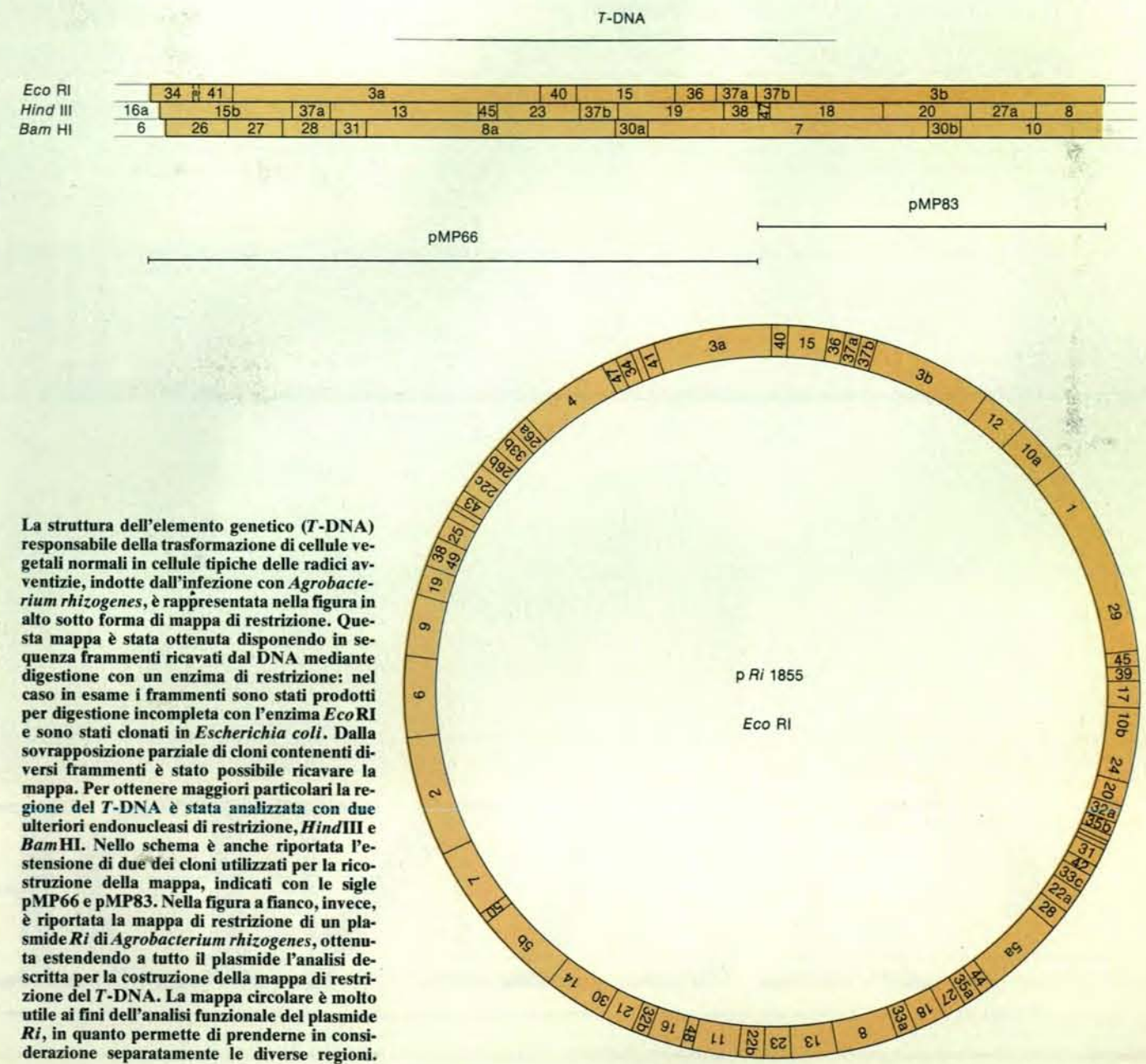
È probabile che negli agrobatteri il complesso sistema di colonizzazione genetica si sia evoluto per gradi: la capacità di indurre la sintesi di opine e quella di causare proliferazione cellulare potrebbero essere acquisizioni distinte, dovute a scambi di materiale genetico con diversi organismi superiori. L'octopina, la prima opina identificata nelle galle del colletto, è infatti un composto assente dal mondo vegetale, ma presente nel mondo animale e comune nel polpo (*Octopus*), da cui è stata inizialmente isolata e da cui ha tratto il nome. D'altra parte la capacità di controllare la crescita e il differenziamento delle cellule vegetali farebbe pensare a un'origine vegetale di quest'altra funzione del *T*-DNA.

Si tratta per il momento solo di stimolanti ipotesi, ma quello che è certo è che il *T*-DNA degli agrobatteri costituisce un formidabile strumento oltre che per la manipolazione genetica delle piante superiori anche per la comprensione a livello molecolare di alcuni dei loro più fondamentali processi fisiologici.

T-DNA

Eco RI: 34 | 41 | 3a | 40 | 15 | 36 | 37a | 37b | 3b
Hind III: 16a | 15b | 37a | 13 | 45 | 23 | 37b | 19 | 38 | 47 | 18 | 20 | 27a | 8
Bam HI: 6 | 26 | 27 | 28 | 31 | 8a | 30a | 7 | 30b | 10

pMP83

pMP66

La struttura dell'elemento genetico (*T*-DNA) responsabile della trasformazione di cellule vegetali normali in cellule tipiche delle radici avventizie, indotte dall'infezione con *Agrobacterium rhizogenes*, è rappresentata nella figura in alto sotto forma di mappa di restrizione. Questa mappa è stata ottenuta disponendo in sequenza frammenti ricavati dal DNA mediante digestione con un enzima di restrizione: nel caso in esame i frammenti sono stati prodotti per digestione incompleta con l'enzima *Eco*RI e sono stati clonati in *Escherichia coli*. Dalla sovrapposizione parziale di cloni contenenti diversi frammenti è stato possibile ricavare la mappa. Per ottenere maggiori particolari la regione del *T*-DNA è stata analizzata con due ulteriori endonucleasi di restrizione, *Hind*III e *Bam*HI. Nello schema è anche riportata l'estensione di due dei cloni utilizzati per la ricostruzione della mappa, indicati con le sigle pMP66 e pMP83. Nella figura a fianco, invece, è riportata la mappa di restrizione di un plasmide *Ri* di *Agrobacterium rhizogenes*, ottenuta estendendo a tutto il plasmide l'analisi descritta per la costruzione della mappa di restrizione del *T*-DNA. La mappa circolare è molto utile ai fini dell'analisi funzionale del plasmide *Ri*, in quanto permette di prenderne in considerazione separatamente le diverse regioni.

La riproduzione digitale del suono

Rappresentando il suono come una serie di numeri discreti è possibile ottenere una incisione di qualità superiore e, di conseguenza, anche un miglioramento senza precedenti nella riproduzione delle esecuzioni musicali

di John Monforte

Oggi si può ascoltare Enrico Caruso in una registrazione di gran lunga superiore qualitativamente a tutte le incisioni che fece in vita. Questo perché le vecchie registrazioni su cilindro di cera sono state digitalizzate: i suoni musicali sono stati riprodotti come una serie di numeri discreti, anziché come forme d'onda analogiche, cioè variabili con continuità, come era sulle vecchie registrazioni. Il segnale digitalizzato è stato elaborato da un calcolatore per eliminare tutti gli artefatti introdotti dalla tromba acustica di registrazione. In altri termini, il suono della voce di Caruso è stato migliorato. Si è trattato, in effetti, di un esperimento e il processo ha bisogno di ulteriori perfezionamenti; tuttavia oggi è possibile apprezzare tutta la grandezza di questa voce tenorile.

Può darsi che la digitalizzazione serva a migliorare le registrazioni di altri grandi artisti del passato, ma la sua funzione principale è quella di permettere una elevata qualità di incisione di esecuzioni musicali contemporanee.

Il processo digitale costituisce un modo del tutto nuovo per memorizzare, elaborare e anche per generare segnali audio. Il fonografo è il primo componente dei sistemi audio che possa avvantaggiarsene, ma con il tempo ogni componente della catena, dalla registrazione alla riproduzione, sarà sostituito da una controparte che tratterà i suoni sotto forma di dati digitali. L'elaborazione digitale può essere utilizzata anche per comporre e riprodurre musica con l'uso di suoni che non hanno un corrispondente nello strumentario tradizionale del musicista.

Una tipica incisione digitale potrebbe arrivare al consumatore sotto forma di un disco alluminiato di 120 millimetri di diametro. Il suono digitalizzato è conservato come una serie di cavità e di aree piane, di dimensioni microscopiche. Il dispositivo di riproduzione, che va a sostituire il fonorivelatore (*pick-up*) o lo stilo dei precedenti giradischi, è un fascio laser. Il suono è portato poi fino all'ascoltatore da amplificatori e altoparlanti.

La qualità sonora più elevata di queste registrazioni è dovuta in parte al fatto che la digitalizzazione permette di superare molte limitazioni della registrazione analogica. In una registrazione analogica il segnale (musica o qualche altro tipo di suono) può essere descritto come una variazione continua di ampiezza nel tempo. Riportato come una curva in un sistema di coordinate ortogonali, il segnale è una serie di onde irregolari. Ogni dispositivo di registrazione accetta una forma d'onda di tal fatta e agisce su di essa in un modo definito dalla funzione di trasferimento del dispositivo stesso (una descrizione matematica del modo in cui il dispositivo riproduce il suono, compresi rumori e distorsioni caratteristici che il dispositivo impone al segnale che lo attraversa). Alla fine il dispositivo emette una forma d'onda di uscita elettroacustica, che è una replica della forma d'onda di ingresso.

La funzione di trasferimento si applica a dispositivi come gli amplificatori (che si limitano ad aumentare l'ampiezza del segnale) o ai trasduttori (che convertono un segnale da una forma di energia a un'altra: un altoparlante, per esempio, converte energia elettrica in energia acustica). La funzione di trasferimento deve essere il più lineare possibile, cioè deve seguire con la massima esattezza possibile la forma d'onda che rappresenta il suono di partenza. Qualunque deviazione dalla linearità imprime una distorsione alla forma d'onda. Tutti i sistemi analogici sono, almeno in qualche misura, non lineari.

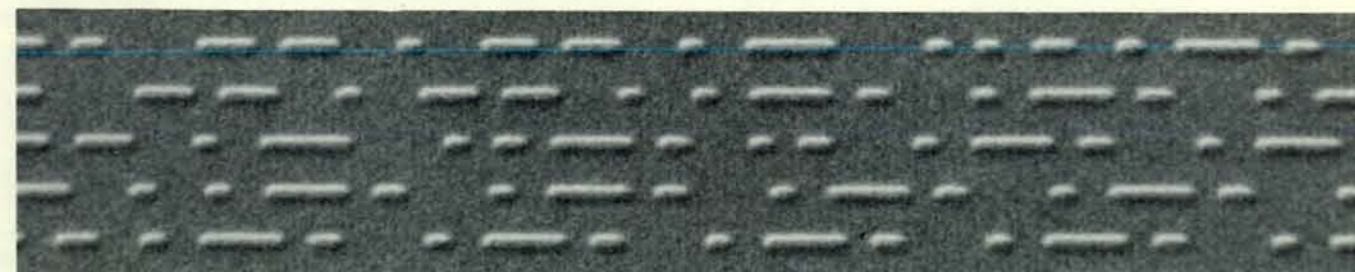
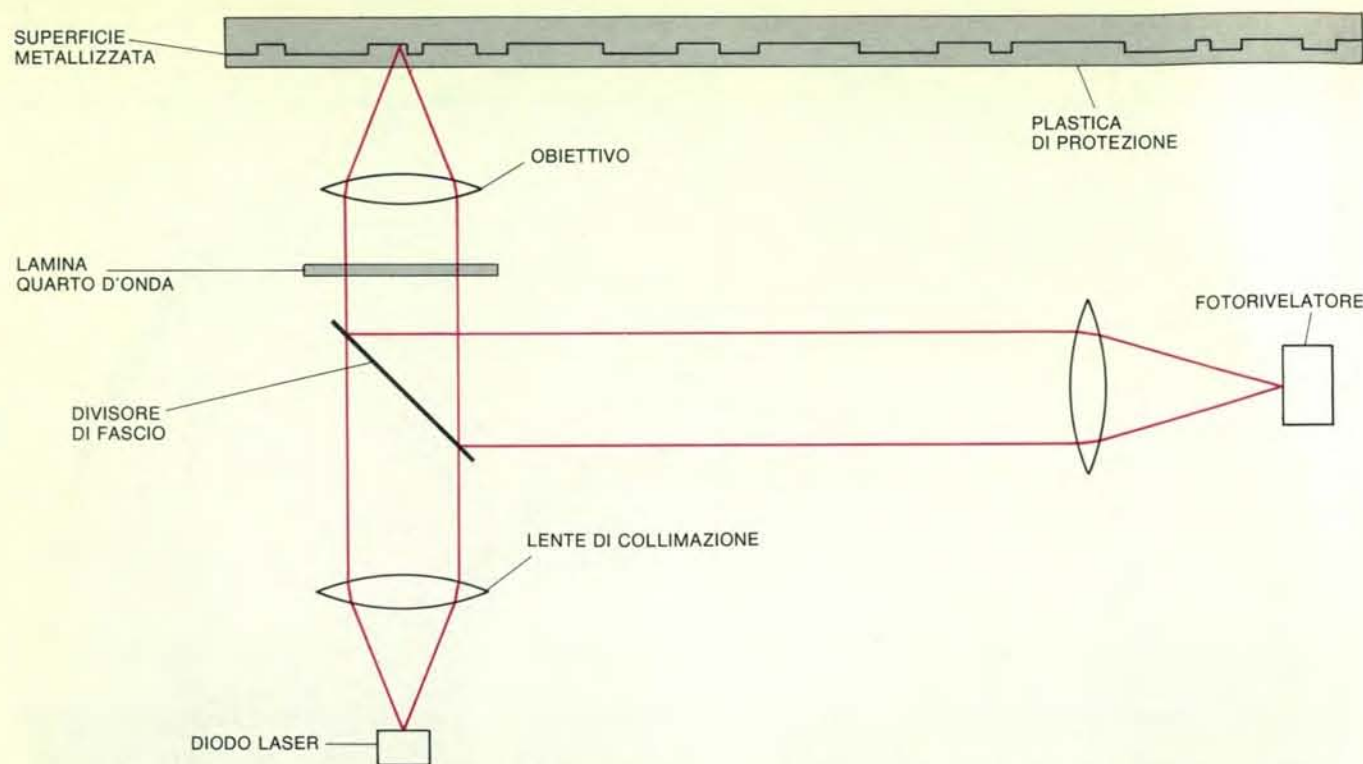
Il problema del rumore

Anche il rumore è un problema onnipresente, per i sistemi analogici. Quando un segnale viene trasformato da un trasduttore (un microfono, per esempio), il livello di uscita è molto vicino al livello del rumore termico causato da vibrazioni casuali degli elettroni nel trasduttore. Questa sfortunata circostanza è dovuta al fatto che il rendimento di un trasduttore si aggira attorno all'1 per cento. Quando il segnale è memorizzato in un supporto di registrazione, il supporto stesso diventa una sorgente di rumore, a causa della sua natura «granulare». Questa granularità deriva da fattori come le dimensioni dei domini magnetici in un nastro da registrazione o l'irregolarità della superficie di un disco. Il rumore fissa un limite inferiore al potere di risoluzione del supporto di memorizzazione.

Tutti i componenti di un sistema audio hanno anche un limite superiore. Le dimensioni del segnale possono crescere solo fintantoché il sistema non è sovraccarico. Questa limitazione è universale e si può trovare in un amplificatore che non può fornire ulteriore potenza, in un altoparlante in cui l'elemento di produzione del suono raggiunge i limiti di escursione o in un solco di un disco che va a sovrapporsi a un solco adiacente. Una misura della fedeltà (ma non l'unica) è la differenza fra il massimo segnale possibile e il livello di rumore, differenza che viene definita «gamma dinamica» e che viene misurata in decibel.

Un'altra limitazione fondamentale presente in tutti i dispositivi di registrazione è la larghezza di banda. Può essere descritta come un limite imposto alla velocità con cui il sistema può rispondere alle rapide variazioni di ampiezza caratteristiche dei suoni. La causa può essere inerzia meccanica, fattori elettrici o tutta una serie di altre fonti: in ogni caso viene imposta una limitazione, che definisce la frequenza più elevata che può transitare attraverso il dispositivo.

Tutti i suoni mettono a dura prova la larghezza di banda (cioè l'intervallo di frequenze passanti) di un sistema audio. Un suono è formato da una frequenza (la fondamentale), che permette all'ascoltatore di percepire l'altezza del suono, e da una serie di componenti a frequenza più elevata (chiamate componenti parziali o semplicemente parziali). Le parziali determinano il timbro del suono. Tutti questi elementi esistono simultaneamente. Nei suoni musicali meno dissonanti le parziali sono disposte in una serie di frequenze che sono multipli interi (o all'incirca)

Il prodotto comunemente in commercio per la riproduzione digitale del suono è il Compact Disc, frutto di una collaborazione tra la Philips Corporation e la Sony Corporation. Sul disco, che ha un diametro di 120 millimetri e uno spessore di 1,2 millimetri, i dati digitali sono memorizzati sotto forma di rilievi microscopici e di aree piane, che rappresentano rispettivamente le cifre binarie 1 e 0. Un 1 è l'inizio o la fine di un rilievo; uno 0 è un'area senza variazioni. Un fascio laser svolge la funzione che ha in un giradischi tradizionale il fonorivelatore (*pick-up*). Avendo una capacità di 500 megabyte, questo tipo di disco è la forma di supporto a densità massima di registrazione di dati.

La superficie di un Compact Disc è caratterizzata da piccoli rilievi (alti 0,16 e larghi 0,6 micrometri), che rappresentano i dati memorizzati. Il segnale registrato nella sezione di disco riprodotta è il silenzio e di conseguenza il suo andamento è molto ripetitivo, interrotto solo dalle parole digitali che servono a correggere errori o a delimitare gruppi di dati. La microfotografia mostra la faccia del disco letta dal fascio laser.

Il meccanismo di lettura a laser trasforma l'incisione del disco in variazioni nell'intensità di un fascio di luce monocromatico collimato. Il fascio, che è generato da un diodo laser, passa attraverso una lente di collimazione che corregge la sua tendenza a divergere. Un divisore di fascio a polarizzazione separa il fascio concentrato, inviandone parte a un fotorivelatore (*a destra*); il resto del fascio passa attraverso un altro filtro che ne fa ruotare di 90 gradi l'asse di polarizzazione. Un obiettivo focalizza il fascio sulla superficie del disco. I rilievi presenti sul disco, complementari alle cavità incise nel processo di registrazione della matrice, hanno dimensioni paragonabili alla lunghezza d'onda della luce incidente (qui i rilievi sono esagerati per motivi di chiarezza) e di conseguenza diffondono la luce, impedendole di ritornare al rivelatore il quale registra questa diminuzione di intensità del fascio. Le variazioni di intensità sono lette dal sistema come sequenze di cifre binarie.

della fondamentale. Perché un sistema audio sia davvero utile, è preferibile metterlo in grado di rispondere a tutto l'intervallo di frequenze udibili da parte dell'orecchio umano: una larghezza di banda, cioè, di circa 20 000 hertz.

Il processo di digitalizzazione è stato realizzato ai Bell Telephone Laboratories negli anni venti, con l'obiettivo di superare le limitazioni della registrazione analogica. Solo da poco, però, sviluppi tecnologici come l'elaborazione elettronica e microchip di grande versatilità hanno consentito un'evoluzione della tecnica che ha reso i sistemi digitali davvero superiori qualitativamente (e concorrenziali nei costi) rispetto ai sistemi analogici. L'evoluzione sicuramente non ha ancora raggiunto l'epilogo.

In un sistema audio digitale la forma d'onda è ridotta a una serie di numeri che descrivono il suo andamento. I numeri poi sono trattati da un calcolatore, che può memorizzarli in vista di un recupero in tempi successivi oppure può elaborarli seguendo algoritmi matematici per modificare, a fini acustici o musicali, le forme d'onda che rappresentano. I segnali digitali che rappresentano i numeri possono accumulare rumore e distorsione, quando vengono memorizzati o elaborati; tuttavia, finché non vanno perduti e restano ragionevolmente riconoscibili, permettono di ricostruire la forma d'onda originale con scarse perdite qualitative.

Il primo passo nella registrazione digitale è la digitalizzazione di un segnale audio. Il dispositivo che svolge questa funzione prende il nome di convertitore analogico-digitale o ADC (da *analogue-to-digital converter*). La forma d'onda di un segnale analogico è una funzione bidimensionale (dà l'ampiezza del segnale in funzione del tempo), e pertanto si deve predisporre un segnale digitale che descriva con precisione ambedue le dimensioni.

In primo luogo, il segnale viene campionato a intervalli regolari determinati da una frequenza di campionamento standardizzata; i dati poi sono raccolti a una velocità predeterminata. Quando viene il momento di ricostruire i dati sotto forma di segnale analogico, l'informazione deve essere restituita esattamente alla stessa frequenza per evitare variazioni di velocità o di altezza del suono.

La tensione che rappresenta l'ampiezza del segnale al momento della conversione viene poi memorizzata in modo analogico per un tempo abbastanza lungo da consentirne la misurazione e l'assegnazione di un numero che rappresenta il valore di tensione. Questo passo prende il nome di quantizzazione. Ogni campione genera un nuovo numero: così si forma un flusso di dati che può essere elaborato da qualunque sistema digitale.

Al fine di ricostruire il segnale, i dati sono presentati a un convertitore digitale-analogico o DAC (da *digital-to-analogue converter*) alla frequenza di campionamento. Il convertitore genera una tensione che rappresenta il valore dei dati e conserva la tensione fino a che non riceve una nuova parola di dati. Il segnale risultante è una replica della forma d'onda originale.

Fonti di errore

A causa di un campionamento inadeguato il segnale può essere distorto in due modi distinti, legati l'uno alla frequenza di campionamento e l'altro alla quantizzazione. Ciascuno dei due problemi può essere risolto mediante un campionamento più fine, ma a un costo: è necessaria una maggiore capacità di gestione dei dati. Anche così, non c'è modo di evitare completamente queste distorsioni, perché la forma d'onda analogica è una funzione continua sia rispetto all'ampiezza, sia rispetto al tempo, e qualunque rappresentazione digitale invece deve essere discreta. Una volta digitalizzato, tuttavia, il segnale può essere trasmesso o copiato indefinitamente senza accumulo di distorsioni o di rumore oltre a quelli presenti al momento della digitalizzazione.

Il primo modo in cui un segnale può essere distorto è determinato da una frequenza troppo bassa di campionamento. Se questo non è abbastanza frequente, l'ampiezza del segnale può variare di una quantità superiore a un quanto nell'intervallo tra due campionamenti successivi. (Un quanto è la dimensione del «gradino» più piccolo usato per descrivere il segnale.) In effetti il segnale cambia troppo rapidamente perché il sistema possa registrarlo con precisione. Quando il segnale viene ricostruito, questo cambiamento viene omesso e si perde informazione.

Variazioni di ampiezza così rapide sono causate da componenti del suono di frequenza superiore alla metà della frequenza di campionamento. La presenza di queste componenti fa sì che il segnale ricostruito sembri contenere una componente di frequenza pari alla frequenza di campionamento meno la frequenza originale. Questo fenomeno di disturbo prende il nome di «*aliasing*» o «*foldover*». Le componenti del segnale sottocampionate sono riprodotte come componenti del tutto nuove, in un rapporto non armonico con la fondamentale.

Il suono dovuto all'*aliasing* è non musicale. In effetti, è molto strano. Per evitare questo risultato il campionamento deve avvenire abbastanza spesso da includere tutte le parziali importanti del segnale. È essenziale inoltre disporre un filtro prima del campionatore per limitare forzatamente la larghezza di banda ed eliminare tutte le componenti indesiderate. Il filtraggio va fatto nel campo analogico, prima di qualsiasi digitalizzazione.

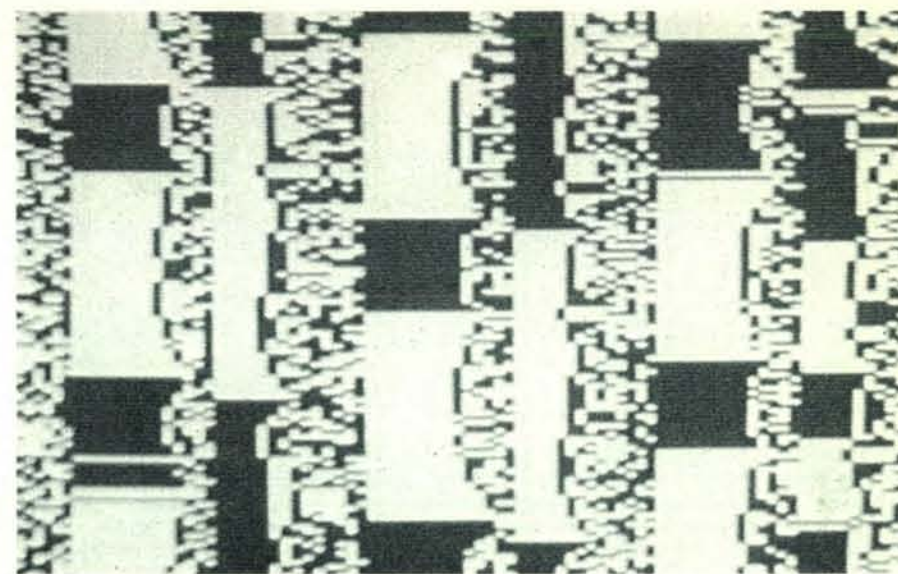
Infine, il segnale deve essere filtrato dopo la ricostruzione. Il segnale replicato ha infatti un caratteristico andamento «a gradini» causato dal processo di campionamento discreto. Questi spigoli netti rappresentano componenti di alta frequenza non presenti nel segnale originale.

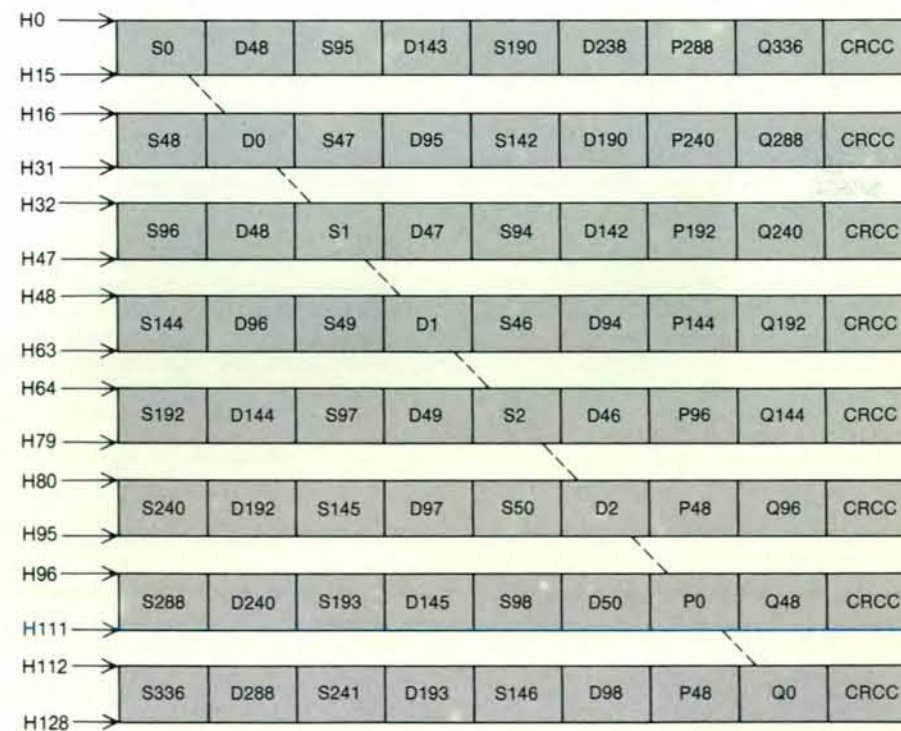
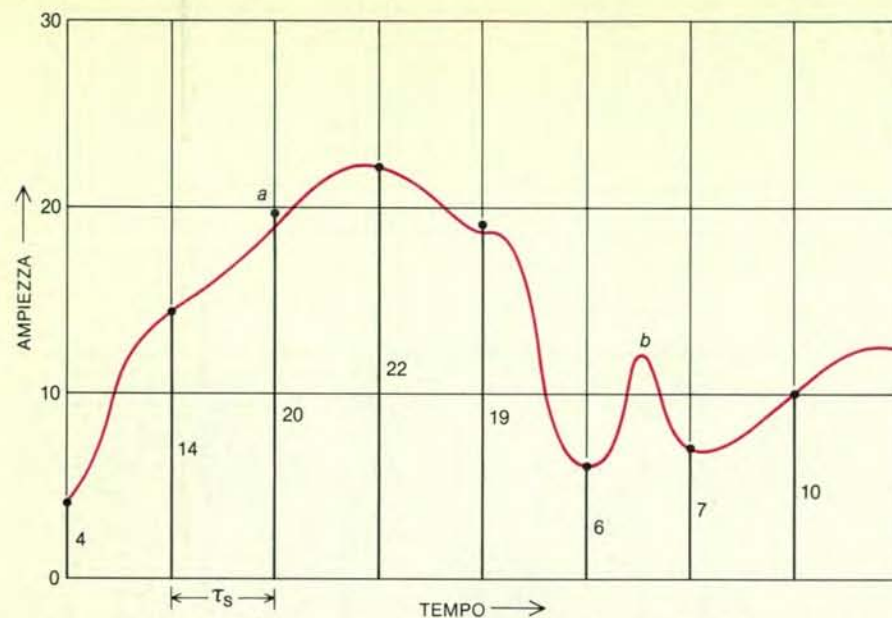
Errore di quantizzazione

La seconda forma di distorsione prende il nome di errore di quantizzazione e sorge quando la parola digitale non è abbastanza grande da descrivere l'ampiezza con sufficiente dettaglio. Quando il livello di ampiezza si trova fra due quanti adiacenti, il digitalizzatore è costretto a scegliere fra i due, causando un errore.

Sul lungo periodo questi errori sono casuali. Il suono è analogo al rumore, se non per il fatto che capita solo in presenza di un segnale. Quello che si sente è un ronzio che sembra in rapporto con il segnale. La soluzione sta nel miscelare rumore analogico di una certa ampiezza con il segnale in ingresso per mascherare il rumore di quantizzazione, dopodiché il rumore suona casuale e la percezione di distorsione si riduce sostanzialmente.

Una volta rappresentata con un numero l'ampiezza del segnale a un dato istante, questo valore deve essere memorizzato. Poiché i dati sono trattati da un calcolatore digitale, viene formata una parola binaria. Il sistema più comunemente impiegato prende il nome di modulazione a codice di impulsi lineare, un'idea sviluppata nel 1939 da A. H. Reeves dei Bell Laboratories.

Il numero viene rappresentato tipicamente sotto forma di una parola di 16 bit. I quanti (ossia i «gradini» della rappresentazione digitale) hanno tutti la stessa dimensione. Il valore pari alla metà del massimo (1000000000000000) rappresenta la «condizione di assenza di segnale», cioè il silenzio. Il numero cresce, tendendo al numero formato da tutte cifre 1, quando l'onda si muove in una direzione (positiva) rispetto alla posizione di riposo, mentre diminuisce tendendo al numero formato da sole cifre 0, quando l'onda si muove in direzione opposta (negativa). Il sistema lineare a 16 bit è incorporato nella maggior parte dei dischi digitali in commercio e permette una gamma dinamica di 96 decibel, superiore quindi di circa 20 decibel alle possibilità di un registratore analogico e di circa 30 decibel a quelle dei dischi fonografici.

La memorizzazione digitale su nastro per videoregistrazione genera su uno schermo televisivo una immagine come questa, nella quale i punti bianchi e neri rappresentano le cifre binarie.

H0 – H15	S0	D48	S95	D143	S190	D238	P288	Q336	CRCC
H16 – H31	S48	D0	S47	D95	S142	D190	P240	Q288	CRCC
H32 – H47	S96	D48	S1	D47	S94	D142	P192	Q240	CRCC
H48 – H63	S144	D96	S49	D1	S46	D94	P144	Q192	CRCC
H64 – H79	S192	D144	S97	D49	S2	D46	P96	Q144	CRCC
H80 – H95	S240	D192	S145	D97	S50	D2	P48	Q96	CRCC
H96 – H111	S288	D240	S193	D145	S98	D50	P0	Q48	CRCC
H112 – H128	S336	D288	S241	D193	S146	D98	P48	Q0	CRCC

La struttura di un «fotogramma» di dati digitali registrato su nastro per videoregistrazione (un'alternativa al disco) è costituita da 245 righe simili alle 128 rappresentate qui. Ogni gruppo di 16 righe contiene tre parole di dati provenienti dal canale sinistro (*S*) del nastro, tre dal canale destro (*D*), due parole associate con il metodo di correzione degli errori chiamato controllo di parità (*P*, *Q*) e una parola di codice per la tecnica di correzione degli errori chiamata codice di controllo di ridondanza ciclica (*CRCC*). Poiché il nastro tende a perdere in un sol colpo intere righe orizzontali, i dati sono «interfogliati» (*linea in diagonale*). Le parole per la correzione degli errori così sparpagliate possono essere utilizzate per ricostruire qualunque dato che si perda.

Il processo di digitalizzazione comincia quando un convertitore analogico-digitale riceve una forma d'onda analogica (*in colore*). La forma d'onda è costituita da una tensione che varia con continuità nel tempo. Il convertitore effettua il campionamento della forma d'onda circa 48 000 volte al secondo (τ_s è l'intervallo di tempo fra un campionamento e il successivo). Ogni campione è rappresentato da un pallino. L'ampiezza dei suoni campionati è indicata con numeri decimali. Fra le fonti di imprecisione vi sono graduazioni di campionamento troppo ampie (*a*). Qui il valore vero è minore di 20 ma maggiore di 19. Il sistema, che non è in grado di risolvere differenze così piccole, assegna arbitrariamente il valore più vicino; una riproduzione conterrà pertanto questo «errore di quantizzazione». Un'altra imprecisione (*b*) può presentarsi quando il segnale subisce una variazione in un tempo inferiore all'intervallo di campionamento (cioè in meno di un «quanto» dell'asse temporale). Una variazione di questo tipo non verrà osservata perché la frequenza di campionamento è troppo bassa. Quanti più piccoli e un campionamento più frequente possono ridurre questi errori, generando però parole di dati più grandi e in maggior numero cosicché il sistema deve essere in grado di trattare una maggior quantità di dati.

Benché la modulazione a codice d'impulsi lineare costituisca il metodo di conversione più comune, senza dubbio non è il migliore per tutte le applicazioni. Sono disponibili altri metodi di conversione. Uno di questi, che consente anche di ridurre la quantità di informazione necessaria per rappresentare il segnale, consiste nel memorizzare solamente la differenza fra un campione dato e il precedente. Questa tecnica prende il nome di modulazione delta (per «differenza»).

La modulazione delta si basa sull'ipotesi che il segnale campionato sia prevalentemente un segnale continuo e che le variazioni più ampie abbiano luogo su periodi di tempo più lunghi. L'ipotesi risulta valida in particolar modo nel caso della musica, poiché l'ampiezza delle parziali superiori degli strumenti musicali è generalmente minore dell'ampiezza della fondamentale.

Vi potrà capitare di sentir discutere su quale sia il metodo di codificazione che porta ai suoni più realistici. Il metodo lineare è il più semplice. Se il segnale è campionato abbastanza spesso e viene rappresentato da un numero abbastanza grande di bit, si può ottenere un grado di fedeltà molto elevato.

Il metodo lineare, invece, non è necessariamente il più efficiente per quanto riguarda l'utilizzazione dello spazio di memoria. Il cervello tende a ignorare i segnali sensoriali che permangono a lungo e a porre maggiore attenzione alle variazioni. (Fate caso alle scarpe quando le mettete o quando le levate, ma ben difficilmente, in genere, nell'arco della giornata.) Si può sostenere allora che la modulazione delta sfrutti lo spazio di memoria nel modo più efficiente. Lo spazio che si recupera grazie all'immagazzinamento di parole più piccole può essere utilizzato per un campionamento più frequente, la qual cosa a sua volta rende meno stringenti i requisiti del filtraggio e riduce le possibilità di sovraccarico dovute a segnali che variano rapidamente.

La memorizzazione dei dati

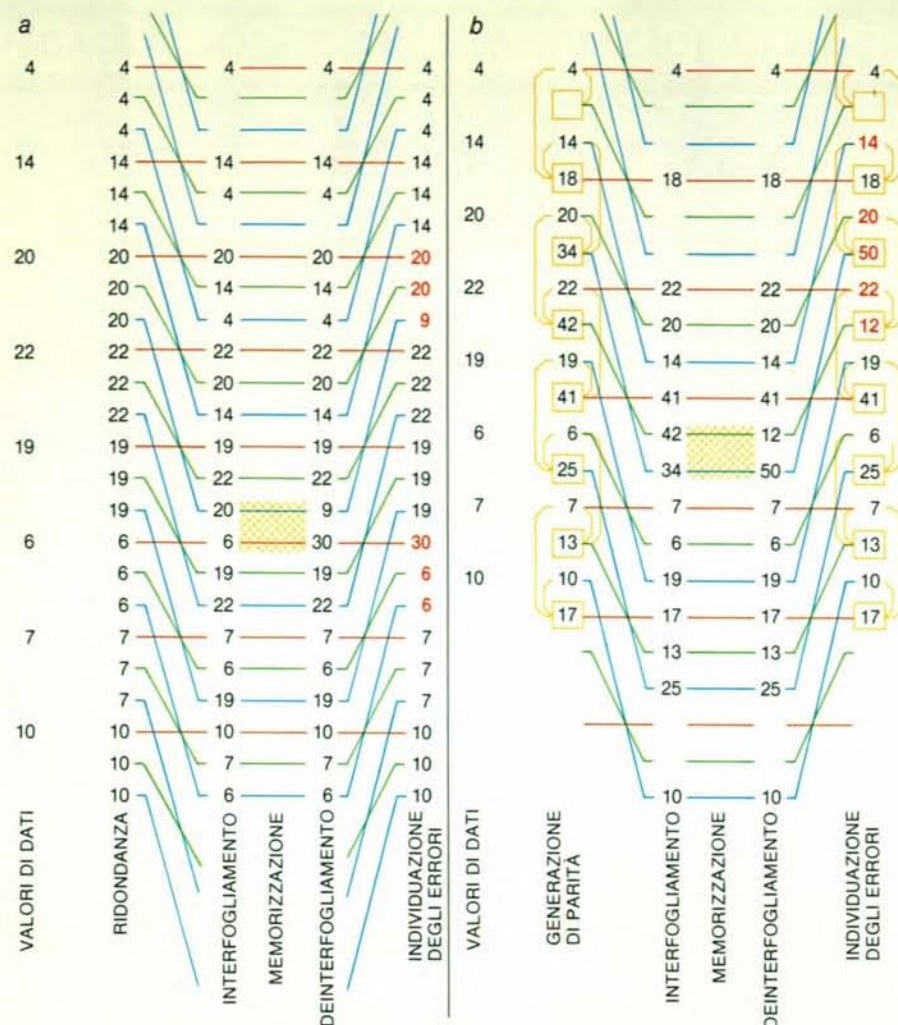
Vi sono vari modi per memorizzare i dati ottenuti dal campionamento. Il più semplice e il più diretto consiste nell'inserire le informazioni negli stessi sistemi di memoria utilizzati da un calcolatore digitale. Tuttavia il grande volume di dati rende poco pratico questo tipo di memoria. Un segnale digitale campionato a una frequenza di 40 chilohertz riempirebbe oltre 200 megabyte di memoria nel giro di un minuto. Inoltre, la capacità di accesso casuale tipica delle memorie dei calcolatori non è necessaria per la musica, che è per sua natura sequenziale.

Per i calcolatori digitali, il metodo più comune di memorizzazione sequenziale è il nastro magnetico. Molti registratori audio sfruttano la tecnologia (e a volte anche i componenti) delle unità a nastro magnetico dei calcolatori digitali. I registratori audio sono diversi in quanto ricevono contemporaneamente più flussi di dati (da vari trasduttori), che debbono essere immagazzinati in tracce adiacenti. In questo modo il tecnico può registrare le singole parti musicali in momenti diversi e combinarle in seguito.

Un'altra tecnica molto diffusa consiste nella memorizzazione dei dati digitali su un nastro magnetico per videoregistrazione. Un videoregistratore ha una larghezza di banda necessariamente grande per poter alloggiare l'enorme quantità di informazioni necessaria per descrivere tutta la gamma di luminosità di un'immagine video. Nella registrazione audio ogni bit è rappresentato da un singolo pixel (elemento d'immagine), che è bianco o nero a seconda che il bit sia, rispettivamente, 1 o 0. Il segnale digitale viene trasformato in un comune segnale video, cosicché è possibile utilizzare un videoregistratore qualsiasi.

Per una diffusione di massa, perché cioè le registrazioni digitali possano arrivare fino al piccolo consumatore, il nastro magnetico non rappresenta un supporto ideale, perché ciascuna copia deve essere ottenuta dalla matrice (il «*master*») con il trasferimento sequenziale dei suoni registrati. L'efficienza produttiva richiede un supporto analogo al disco fonografico, che permette la stampa di interi programmi in un solo colpo, con un processo di tipo metallurgico.

Il sistema in uso oggi va sotto il nome di «Compact Disc» o CD, ed è il frutto delle ricerche condotte dalla Philips Corporation in collaborazione con la Sony Corporation. Il disco ha un diametro di 120 millimetri e uno spessore di 1,2 millimetri; può memorizzare fino a 74 minuti di programmazione. I dati digitali sono memorizzati originariamente sotto forma di una serie di cavità e di assenze di cavità, che rappresentano rispettivamente le cifre binarie 1 e 0. Si ha un 1 dove inizia o finisce una cavità, uno 0 là dove non vi è variazione. La superficie incisa è fusa in plastica da uno stampo ottenuto da una matrice: le cavità quindi diventano piccoli rilievi. La superficie del disco è alluminiata per conferirle riflettività (il dispositivo di riproduzione è un fascio laser) e protetta da uno strato di plastica trasparente.

Quando il disco viene riprodotto, il fascio laser va a colpire o un'area piana o un rilievo. Se incide su un'area piana, viene riflesso verso un fotorivelatore. Se incide su un rilievo, viene diffuso. Il sensore registra una variazione di intensità come un 1, la mancanza di variazione come uno 0. È evidente che la superficie del disco deve rimanere sufficientemente piana e il sistema ottico sufficientemente stabile perché la distanza fra l'obiettivo del sistema ottico del laser e il disco non cambi. In caso contrario il fascio non sarà più ben focalizzato e andrà a investire un'area maggiore di quella che corrisponde a un singolo bit. Il meccanismo ottico in effetti

Il controllo degli errori può essere ottenuto in vari modi. Qui sono stati applicati due metodi diversi ai dati ottenuti nel campione mostrato nell'illustrazione di pagina 56. Il primo è quello della ridondanza (*a*). Questo metodo richiede che le parole di dati, nella colonna a sinistra, vengano rese ridondanti (*seconda colonna*) mediante ripetizione. Il supporto di memoria però tende a perdere i dati a raffiche; quindi non è saggio immagazzinare le parole ridondanti vicino ai dati originali. I dati pertanto vengono interfogliati secondo uno schema predeterminato, indicato dalle linee in colore. L'interfogliamento viene effettuato prima della memorizzazione (*terza colonna*). Il fatto che siano stati introdotti errori (*area tratteggiata*), così che 20 viene letto 9 e 6 viene letto 30, risulta evidente allorché viene effettuato il processo inverso all'interfogliamento (*quarta colonna*), come indicato dai numeri in colore (*quinta colonna*). Nel secondo metodo (*b*) agli stessi dati sono attribuite parole di controllo di ridondanza ciclica. Qui le parole, racchiuse in rettangoli, sono generate dalla somma di ciascuna coppia di parole di dati. Gli errori sono identificati in fase di recupero quando le somme non corrispondono. Negli esempi 14 + 20 viene letto 50 e 20 + 22 viene letto 12. Poiché ciascuna parola di dati contribuisce a due parole di controllo, è possibile determinare quali siano i valori errati e stabilire così quelli corretti.

PRIMA DELLA MEMORIZZAZIONE:

PAROLA DI DATI #1 19→	010011
+ PAROLA DI DATI #2 6→+	000110
PAROLA DI PARITÀ	010101 = "21"

DOPO IL RECUPERO:

PAROLA #1 MANCANTE	PAROLA #2 MANCANTE	PAROLA DI PARITÀ MANCANTE
PAROLA DI PARITÀ 21→010101	PAROLA DI PARITÀ 010101	PAROLA DI DATI #1 010011
− PAROLA DI DATI #2 6→000110	− PAROLA DI DATI #1 010011	+ PAROLA DI DATI #2 000110
PAROLA DI DATI #1 010011	PAROLA DI DATI #2 000110	PAROLA DI PARITÀ 010101

Il controllo di parità è un altro sistema di protezione contro gli errori. Se si sommano due parole di dati, i cui valori sono 19 e 6, senza il normale riporto della somma binaria, il risultato è una parola di controllo di parità con valore 21. Se dopo il recupero dei dati registrati una delle tre parole contiene un errore o è stata cancellata, si può ricalcolarla (*in basso*) se le altre sono intatte.

è munito di servosistemi che effettuano continuamente regolazioni fini per mantenere costante quella distanza.

La correzione degli errori

La densità di memorizzazione nei sistemi audio tende a essere molto superiore di quella dei calcolatori comuni. Pertanto sono richiesti mezzi molto raffinati per l'identificazione degli errori e la loro correzione. Dal momento che ciascun sistema di memoria ha i propri modi di errore, il tipo di correzione degli errori chiamato in gioco dipende dal metodo di memorizzazione impiegato. I dati memorizzati mediante la tecnica della modulazione a codice di impulsi sono particolarmente sensibili alle perdite casuali di bit. La vulnerabilità deriva dal fatto che i singoli bit rappresentano ordini di grandezza diversi e di conseguenza unità distintive di informazione. Quanto maggiore è l'ordine di grandezza rappresentato da un bit danneggiato, tanto maggiore è il danno per l'uscita che ne risulta. In un caso del genere, quello che si sente è un suono piuttosto esplosivo.

La correzione globale degli errori viene effettuata attraverso la formazione di nuove parole incluse con le parole di dati originali. Le nuove parole, pertanto, sono suscettibili alle stesse fonti di errore delle parole originali e la propagazione degli errori non può essere eliminata completamente. È necessario un metodo di correzione degli errori che lasci solo una probabilità ragionevolmente piccola di permanenza di un errore.

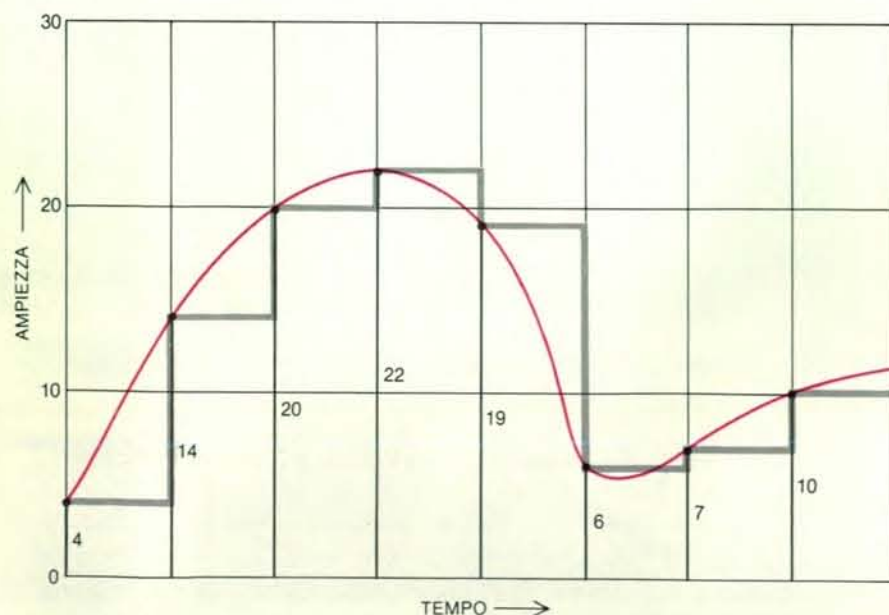
Per poter controllare i dati alla ricerca di errori, l'informazione registrata deve essere immgazzinata temporaneamente in una memoria di transito (*buffer*) dinamica. La memoria di transito deve essere abbastanza ampia da immagazzinare sia la parola di dati sia le parole inserite per il controllo degli errori. Inoltre, gli algoritmi usati per la ricerca e l'eliminazione degli errori debbono essere eseguiti a velocità abbastanza elevata perché i dati corretti possano lasciare la memoria di transito a una velocità accettabile per il sistema di registrazione.

Le strategie fondamentali per la correzione degli errori sono poche e i sistemi di correzione, per la maggior parte, contengono una combinazione di alcune di esse. La più semplice si basa sulla «forza bruta»: ogni parola viene ripetuta una volta. Vi sono alcuni ovvi svantaggi. La quantità di spazio necessaria per la memorizzazione raddoppia; inoltre, non è possibile decidere quale fra le due parole sia corretta senza avanzare ipotesi.

Si può supporre, per esempio, che un'imperfezione nel supporto di registrazione elimini un pezzo di dati alla volta, cosicché gli errori risulteranno concentrati in un'area specifica. Questo fenomeno, che prende il nome di *burst error* o «errore a raffica», è causato comunemente da imperfezioni del supporto. Se nel corso della registrazione le parole e i loro duplicati sono sparpagliati qua e là (procedimento che si definisce *interleaving* o interfoglia-

mento), un esame degli errori può mettere in luce come una parola di ciascuna coppia danneggiata tenda a essere localizzata in prossimità di altre parole che hanno subito danni analoghi. Perché la ridondanza possa essere una tecnica efficace, le parole di controllo debbono essere ampiamente sparpagliate e si debbono esaminare numerose coppie dubbie, prima di poter determinare i valori corretti. Una tecnica del genere richiede una memoria di transito di notevole capacità e un algoritmo assai complesso.

Sono disponibili molte tecniche più eleganti. Un metodo che ha avuto molto seguito comporta la generazione di parole

Il rumore di quantizzazione è la differenza fra il segnale analogico (continuo) reale e la sua rappresentazione digitale (*curva a gradini in colore*). Gli errori non sono mai superiori alle dimensioni di un gradino. La differenza fra il segnale in uscita e il segnale originale costituisce rumore perché la sua ampiezza è piccola e la sua forma d'onda è diversa da quelle dei segnali che l'hanno causata. A differenza del rumore in un sistema analogico, questa forma di rumore è assente quando è assente il segnale. Il risultato è un ronzio che si ha solo in presenza di un suono registrato. Si può mascherare la correlazione sommando un rumore analogico di piccola ampiezza.

Un flusso di dati corretti è inviato a un convertitore digitale-analogico, dove vengono generati, a intervalli di tempo opportuni, i valori analogici rappresentati dai numeri derivati dal campionamento originale, con conseguente produzione del segnale continuo a gradini (*in grigio*). Un filtro elimina l'andamento a gradini, dovuto alla natura discreta dei segnali digitali. In questo modo il segnale in uscita (*in colore*) approssima molto da vicino la forma d'onda di partenza.

di controllo di ridondanza ciclica (fra i professionisti del settore questa tecnica è identificata con la sigla CRC, per *cyclic redundancy check*). Una parola di controllo di questo tipo può essere formata sommando due parole di dati per generare una nuova parola che possiede un bit in più delle due di partenza. Ogni parola di dati viene usata due volte per formare una parola di controllo, ogni volta con una compagna diversa. Se due delle tre parole non subiscono danni, è possibile calcolare la terza. Anche se subiscono danni tutte e tre, è ancora possibile decifrare i valori corretti dalle altre somme formate. Sfortunatamente il numero dei calcoli necessari aumenta di molto, e può non esserci il tempo per effettuarli tutti.

Un altro metodo è il controllo della parità, un metodo comune per la correzione degli errori nel campo dei calcolatori. Si sommano due parole, bit per bit, senza effettuare operazioni di riporto laddove normalmente sarebbe necessario. Il risultato è una nuova parola con lo stesso numero di bit delle due di partenza. Se si trova un errore in una qualunque delle tre parole, le due rimaste inalterate permettono di ricalcolare la terza. I risultati possono essere ottenuti in un solo ciclo di macchina, il che rende il controllo di parità uno strumento molto utile.

Queste e altre tecniche a esse apparentate possono essere combinate in molti modi, ma in ogni caso si deve scendere a qualche compromesso fra la quantità di spazio di memoria necessario, la probabilità che un errore passi inosservato e la velocità a cui gli errori possono essere corretti. Per assicurare la compatibilità, è necessario approntare un metodo standard per ogni dato supporto di memoria.

Ogni tipo di supporto ha i suoi caratteristici modi di errore, e pertanto si può scegliere un metodo diverso per ciascun formato. Il metodo utilizzato per il Compact Disc è in grado di correggere errori a raffica che si estendano su un massimo di 14 000 bit. Gli errori causati praticando nel disco un foro di due millimetri di diametro possono essere corretti con successo, fornendo un'uscita perfetta. Le parole errate che possono passare inosservate sono meno dello 0,005 per cento.

Per ottenere copie prive di errori, il flusso di dati deve essere decodificato, corretto e poi codificato nel formato della copia, anche se questo è identico all'originale. Se non sfuggono errori, la copia suonerà identica all'originale. Inoltre, una matrice consumata può essere «rinfrescata» copiandola. Questo processo può essere ripetuto tutte le volte che si desidera, e ogni generazione sarà indistinguibile dalla prima. Un registratore analogico, invece, a ogni generazione accumula rumore e distorsione.

Talvolta le parole originali e le parole di controllo possono aver subito danni tali da rendere impossibile la correzione di un errore identificato. A questo punto si deve far ricorso a qualche forma di mascheramento dell'errore. Il metodo più semplice (ma anche il più drastico) consiste nel puro annullamento dell'uscita. La maggior parte dei sistemi si comporta in questo modo quando incontra lunghe frasi di dati incomprensibili.

Come rimediare a un errore

La soluzione preferibile è la sintesi di una nuova parola in sostituzione di quella perduta. La falsificazione del materiale è inaccettabile nel mondo dei calcolatori, ma è invece una pratica antica e di larga diffusione fra i musicisti che debbono superare un vuoto di memoria. Un metodo è quello di ripetere la parola precedente. Dato che i segnali musicali restano relativamente costanti da un istante al successivo, questo tipo di mascheramento degli errori è considerato accettabile.

Una pratica più raffinata consiste nell'interpolazione di un valore fra la parola che precede e quella che segue. Questa tecnica richiede un piccolo algoritmo per il calcolo della media: l'algoritmo non è complesso, ma deve svolgere il proprio compito prima che la parola di dati debba lasciare la memoria di transito.

Sono disponibili anche modi più complessi per correggere i dati danneggiati. Uno di questi è l'autocorrelazione, un metodo di intensificazione delle immagini utilizzato per eliminare il rumore dalle immagini digitalizzate inviate a terra dai veicoli spaziali. L'autocorrelazione è uno strumento matematico versatile, che può separare un segnale in due insiemi, uno dei quali contiene le componenti che presentano regolarità o forme distinte (il segnale) e l'altro contiene le componenti casuali (il rumore). Un calcolatore digitale può applicare questo algoritmo o qualunque altro a un segnale audio digitalizzato. Anche se un algoritmo è troppo lungo o complesso per essere eseguito parallelamente a una esecuzione musicale, può comunque svolgere perfettamente la sua funzione per il recupero di una registrazione gravemente danneggiata.

Le possibilità per i compositori

È possibile applicare algoritmi matematici anche per modificare dati privi di danni, in modo da elaborare il segnale per ottenere effetti musicali o sonori. La modifica può assumere la forma di una variazione nel bilanciamento dei bassi o degli acuti, nell'aggiunta di riverbero o nella miscelazione di due o più segnali. È possibile poi aumentare o diminuire la velocità di un segnale senza con questo variarne l'altezza, o viceversa è possibile modificarne l'altezza senza intervenire sulla velocità. In effetti, è possibile elaborare i segnali in modi inimmaginabili, con la possibilità di aprire ai compositori opportunità creative del tutto nuove.

I compositori non avranno solo il vantaggio di nuove forme di elaborazione dei segnali. Con le tecniche digitali si generano oggi suoni utilizzando il calcolatore come un sintetizzatore musicale. La varietà di suoni musicali ed effetti speciali di cui un compositore può disporre con il software progettato per la creazione digitale della musica non ha limiti teorici.

Come uno scorpione del deserto identifica la preda

Una specie notturna di scorpioni, che vive nel deserto del Mojave in California, non vede e non ode gli insetti di cui si nutre, ma porta sulle zampe sensibili recettori che percepiscono il minimo movimento della sabbia

di Philip H. Brownell

Una sera di molti anni fa, Roger D. Farley, dell'Università della California a Riverside, stava osservando con me, su una duna del deserto del Mojave, in California, degli scorpioni che si nutrivano di insetti. A volte, una falena attirata dalle nostre lanterne si posava sulla sabbia vicino a uno scorpione che si voltava fulmineamente e correva nella sua direzione riuscendo spesso a catturarla. Anche leggeri movimenti della sabbia agitata da un fuscello provocavano un attacco vigoroso. Ma l'attenzione dello scorpione non era attratta se una falena tenuta sospesa a mezz'aria a pochi centimetri dallo scorpione batteva le ali: evidentemente lo scorpione non reagiva agli stimoli visivi e acustici, mentre sembrava percepire vibrazioni meccaniche della sabbia, ricavandone le informazioni necessarie a localizzare la preda.

Già a quel tempo avevamo buone ragioni per ritenere che quello scorpione, *Paruroctonus mesaensis,* disponesse di uno straordinario meccanismo di individuazione delle prede. Tanto per cominciare, *Paruroctonus* è un artropodo piuttosto primitivo (gli artropodi sono un phylum che comprende gli insetti, gli aracnidi e i crostacei) e quindi è privo dei sofisticati sensi della vista, dell'udito e dell'olfatto che guidano altri animali verso la preda. In secondo luogo, gli scorpioni sono animali notturni, che escono dalle gallerie solo di notte, per nutrirsi e per accoppiarsi e, dato che la loro capacità visiva è scarsa, è improbabile che la loro efficienza come predatori notturni sia spiegabile con l'informazione che ricevono da questo senso. Infine, una delle più vistose manifestazioni della loro efficienza predatoria è la capacità di individuare prede sepolte, che estraggono con i grossi arti a pinza, i pedipalpi. Certamente nessuno scorpione potrebbe vedere, udire o percepire con l'olfatto un insetto sepolto.

Eppure, l'idea che lo scorpione potesse localizzare la preda percependo le vibrazioni che questa produce nella sabbia era sorprendente. Gli esperti di biologia del comportamento avevano da tempo supposto che gli animali potessero essere messi in allarme da movimenti del substrato, ma senza essere in grado di determinarne la natura e la localizzazione esatta. Questa incapacità è in netto contrasto con l'acuta sensibilità di molti animali alle onde sonore trasmesse attraverso l'acqua o l'aria; in quegli elementi, un animale localizza la sorgente di un segnale rilevando e interpretando differenze minime nei tempi o nell'ampiezza della stimolazione di recettori spazialmente separati. Nei solidi, invece, le perturbazioni meccaniche si trasmettono molto più velocemente e, con lunghezze d'onda maggiori. Ne risulta che le differenze nel tempo di arrivo e nell'ampiezza, a livello di recettori diversi di un singolo animale, sono molto più sottili, troppo sottili - secondo i biologi - per essere captate dal sistema nervoso.

Sembrava che la sabbia fosse il materiale meno adatto a trasmettere informazioni utili perché si pensava che attenuasse molto i movimenti meccanici. Il successo dello scorpione come predatore era quindi un enigma e così decidemmo di studiare nei particolari il suo comportamento, cominciando dalle osservazioni sul campo. In seguito, per spiegare il meccanismo d'individuazione della preda e per accertare le caratteristiche della sabbia che lo rendono attuabile, è stato necessario condurre esperimenti nell'ambiente controllato del laboratorio.

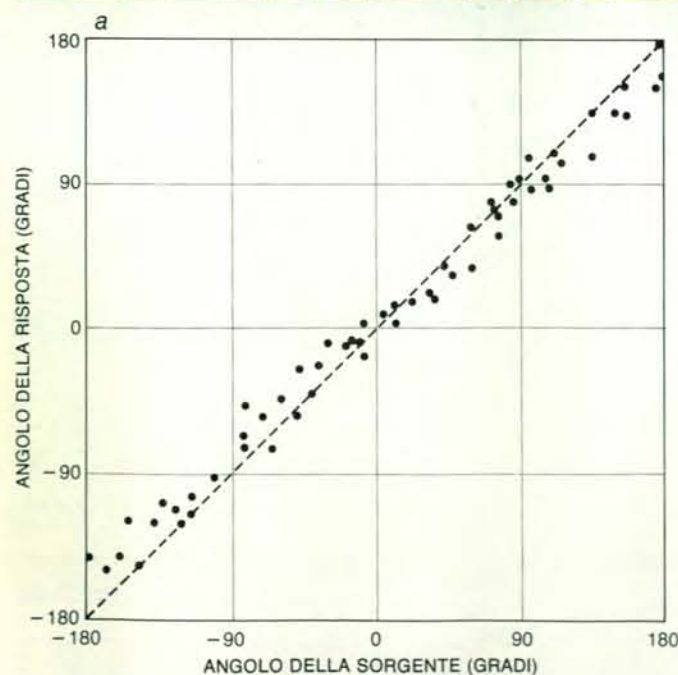
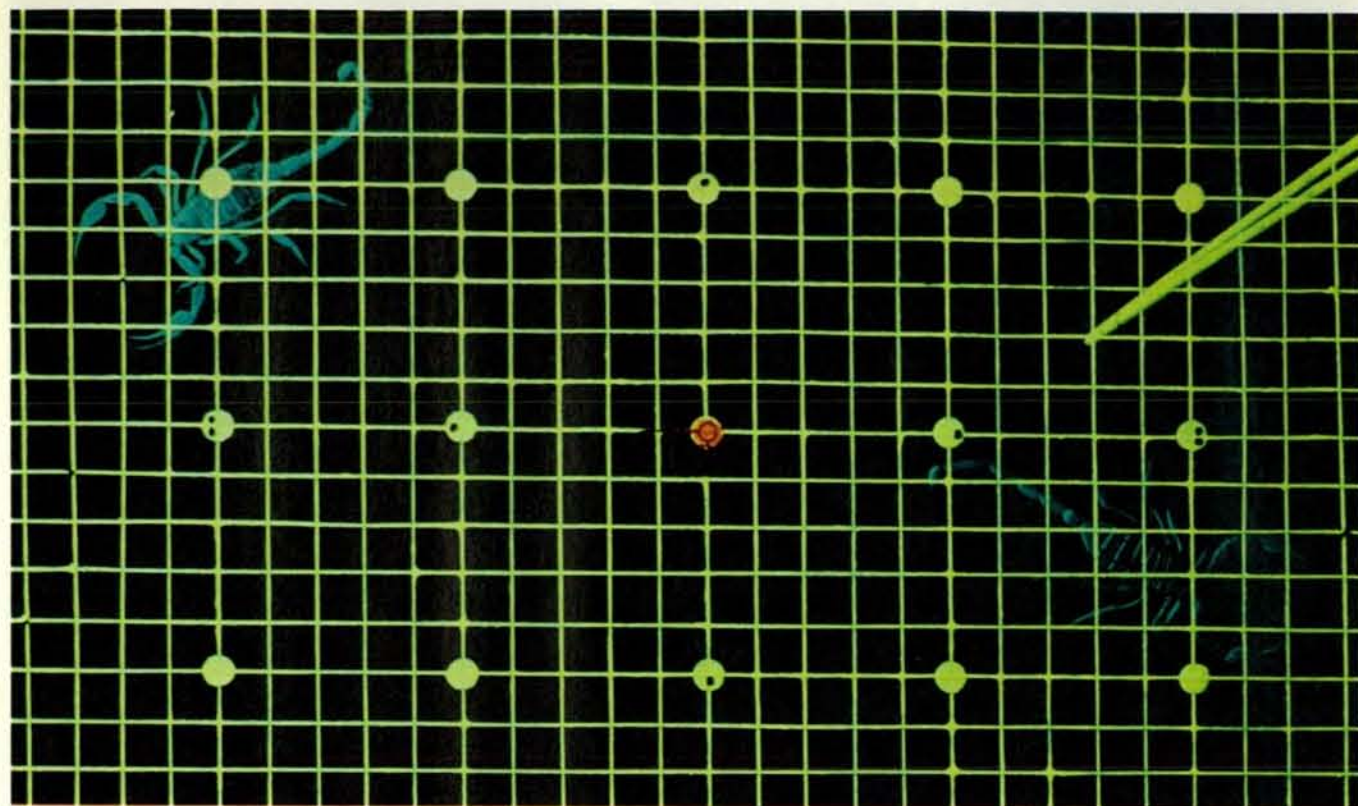
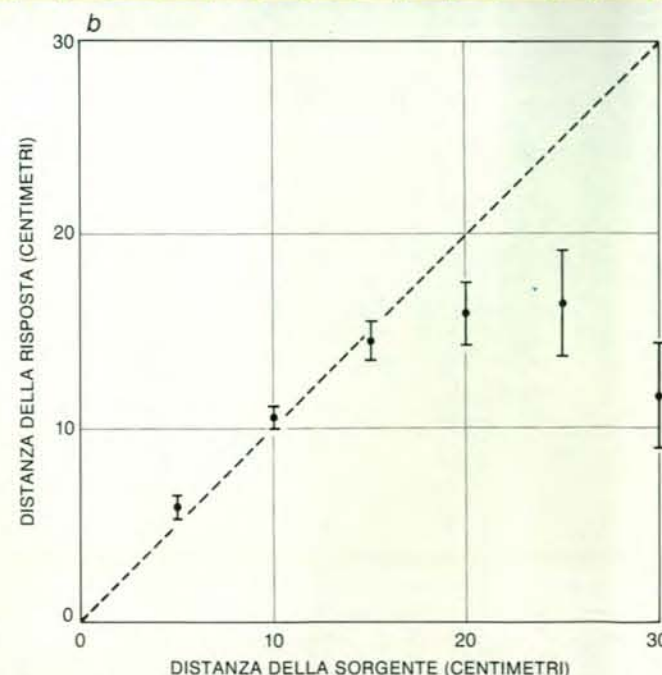
L'habitat di *P. mesaensis,* il deserto del Mojave meridionale, è uno degli ecosistemi più caldi e asciutti dell'America Settentrionale; nelle giornate estive, la temperatura superficiale delle dune supera spesso i 70 gradi centigradi e l'umidità relativa è vicina allo zero. Tuttavia la sabbia assorbe umidità sufficiente alla sopravvivenza di piante cespugliose resistenti alla siccità, come il mesquiete *(Prosopis juliflora)* e molte specie di graminacee annue, le quali servono a stabilizzare le dune contro l'azione del vento e forniscono acqua e cibo a una abbondante varietà di animali.

Gli animali delle dune sono per la maggior parte artropodi notturni; durante il giorno, aracnidi e insetti si difendono dal rischio di essiccamento esistente in superficie scavando gallerie nella sabbia, dove, a una profondità di circa 10 centimetri, la temperatura scende a un livello tollerabile di 40 gradi centigradi e l'umidità supera il 90 per cento. Gary A. Polis della Vanderbilt University ha trovato che *P. mesaensis* è uno dei più comuni artropodi delle dune e anche uno dei più grossi dato che nel corso della sua esistenza, che dura dai cinque ai sei anni, raggiunge una lunghezza di otto centimetri e un peso di quattro grammi, collocandosi così in posizione terminale nella catena alimentare. Di regola, uno scorpione mangia qualsiasi preda che riesca a trattenere per il tempo necessario a paralizzarla con la neurotossina contenuta nel pungiglione che porta all'estremità della coda e quindi molti insetti, compresi i membri più piccoli della sua stessa specie.

Normalmente *P. mesaensis* esce dal nascondiglio per cacciare solo di sera e si dispone all'agguato delle eventuali prede che passeranno a distanza utile; in caso di insuccesso può rimanere immobile per ore, prima di rintanarsi di nuovo per uno o più giorni. Stanley D. Yokota dell'Università dell'Arizona ha dimostrato che uno scorpione può rimediare a una elevata percentuale di insuccessi «disattivando» il proprio metabolismo e rimanendo inattivo nella galleria. Riesce così a sopravvivere per molte settimane e persino mesi, con un solo pasto.

Quando, però, nel suo territorio si presenta una preda adatta, lo scorpione modifica il proprio comportamento in modo appariscente. Dapprima mostra una reazione di allarme: i pedipalpi si aprono e si tendono in avanti mentre lo scorpione assume una posizione eretta al di sopra della sabbia. Qualsiasi movimento suc-

***Paruroctonus mesaensis* occupa una posizione terminale nella catena alimentare delle dune del deserto del Mojave. Alla maturità può raggiungere la lunghezza di otto centimetri e un peso di quattro grammi. Si nutre di insetti e di altri artropodi, che cattura con i grossi arti a pinza (i pedipalpi) e paralizza con la neurotossina contenuta nel pungiglione all'apice della coda. La fotografia, ripresa di notte, mette in evidenza una caratteristica inconsueta della cuticola di questo scorpione notturno, la sua fluorescenza in luce ultravioletta, che consente di osservare il comportamento dell'animale sul campo. (Questa fotografia e quella di pagina 66 sono di Alfred Owczarzak della Oregon State University.)**

cessivo della preda provoca una reazione di orientamento, con cui lo scorpione rivolge rapidamente i pedipalpi verso la preda stessa, spostandosi per diversi centimetri nella sua direzione; se i pedipalpi mancano la presa, lo scorpione si arresta e rimane immobile finché la preda non si muova nuovamente. La sequenza dell'inseguimento dura di solito pochi secondi e in essa si possono verificare da una a cinque risposte d'orientamento, ciascuna preceduta chiaramente da un periodo di immobilità, durante il quale la preda deve smuovere la sabbia perché lo scorpione reagisca.

La fotografia a esposizione multipla, ripresa di notte sul campo in luce ultravioletta, mostra la risposta di uno scorpione che va a caccia di preda quando la sabbia viene lievemente smossa con un bastoncino: esso si volta rapidamente e si dirige verso la sorgente della perturbazione. Se la reazione non lo conduce alla sorgente della perturbazione, lo scorpione rimane immobile in attesa di ulteriori movimenti. Ogni quadratino del reticolo fluorescente sovrapposto misura un centimetro di lato. Con ripetute prove è possibile misurare la precisione dello scorpione nel valutare l'angolo e la distanza della perturbazione che, nel substrato, imita i movimenti della preda. A una distanza compresa tra otto e 10 centimetri, gli animali si rivolgono verso la sorgente della perturbazione con grande precisione. Nella maggior parte delle prove l'angolazione della risposta ha coinciso quasi perfettamente con quella della sorgente *(a)*. Gli scorpioni valutano bene anche distanze fino a 15 centimetri *(b)*. A distanze maggiori, la risposta si allontana in misura significativa dalla precisione assoluta e mostra una maggiore variabilità. (Ogni pallino riportato sul grafico rappresenta la risposta media dell'animale a quella distanza.)

L'abilità dello scorpione nell'incalzare la preda appare in modo particolarmente evidente quando l'animale insegue uno scarafaggio, *Arenivaga investigata*, che vive nel deserto, si nutre di radici e scava gallerie appena al di sotto della superficie sabbiosa. Quando scava nel raggio di mezzo metro da uno scorpione in caccia di preda, quest'ultimo risponde con l'abituale sequenza di movimenti d'orientamento, fino a trovarsi direttamente sopra la preda. A questo punto, ulteriori movimenti dello scarafaggio possono confonderlo ed esso inizia a muoversi in cerchi e a scattare in varie direzioni come se stesse attaccando prede in superficie; ma, nel giro di pochi secondi, modifica il proprio comportamento sollevando l'addome, e affondando ripetutamente i pedipalpi nella sabbia fino a entrare in contatto con lo scarafaggio e ad afferrarlo. Con gli stessi movimenti rotatori di scavo che utilizza per fare una galleria, riporta infine la preda in superficie.

Siccome gli scorpioni sono attivi solo al buio, sembrerebbe difficile valutare quantitativamente il loro raggio d'azione e la loro precisione nel localizzare la preda. Fortunatamente, la loro cuticola presenta una caratteristica singolare: alla luce ultravioletta emette una fluorescenza brillante, giallo-verde, che rende visibile l'animale a molti metri di distanza. Per controllare il comportamento abbiamo disposto sullo stesso cavalletto un proiettore a luce nera e una macchina fotografica, sotto alla quale abbiamo fissato un reticolo fluorescente di riferimento; in questo modo abbiamo potuto sistemare il cavalletto sopra uno scorpione e, in fotografie a esposizioni multiple, abbiamo registrato le risposte d'orientamento dell'animale a movimenti simili a quelli delle prede ottenuti muovendo leggermente la sabbia con un bastoncino.

Dalle nostre misurazioni è risultato che gli scorpioni possono individuare movimenti fino alla distanza di 30 centimetri, mentre a distanze di 10 centimetri o anche meno, la loro valutazione della direzione e della distanza è virtualmente perfetta e la prima risposta al movimento li conduce direttamente al bersaglio. In che modo vi riescono?

Per risolvere il quesito ho deciso di studiare in laboratorio il comportamento di orientamento nello scorpione, ma ciò ha presentato subito un problema in quanto lo scorpione, quando è introdotto in un terrario, non mostra alcuna disposizione alla caccia. Per fortuna, esso ha un altro tipo di comportamento molto simile che si può studiare facilmente: infatti, di norma, uno scorpione fugge le situazioni di pericolo ma, se viene minacciato a sufficienza, solleva l'addome e lo inarca, portando il pungiglione sopra il capo e puntandolo in direzione dello stimolo che ha provocato la sua reazione. Qualsiasi ulteriore perturbazione nel substrato suscita una vigorosa diversione in direzione dello stimolo, come nella risposta predatoria, ma senza movimento in avanti.

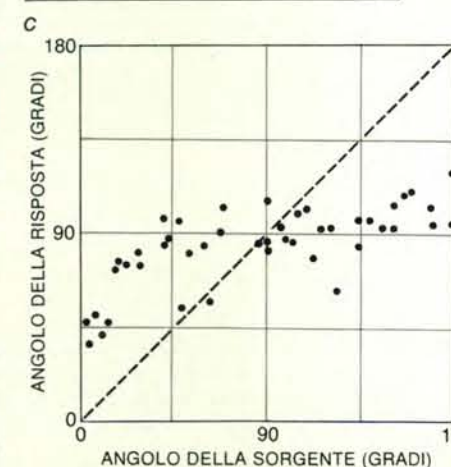
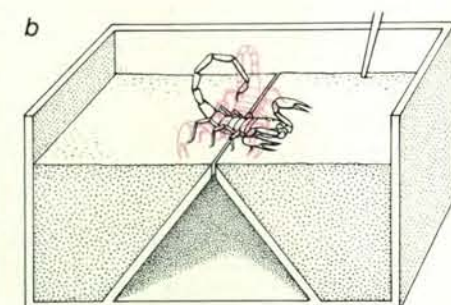
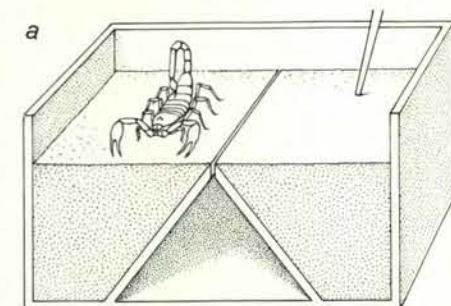
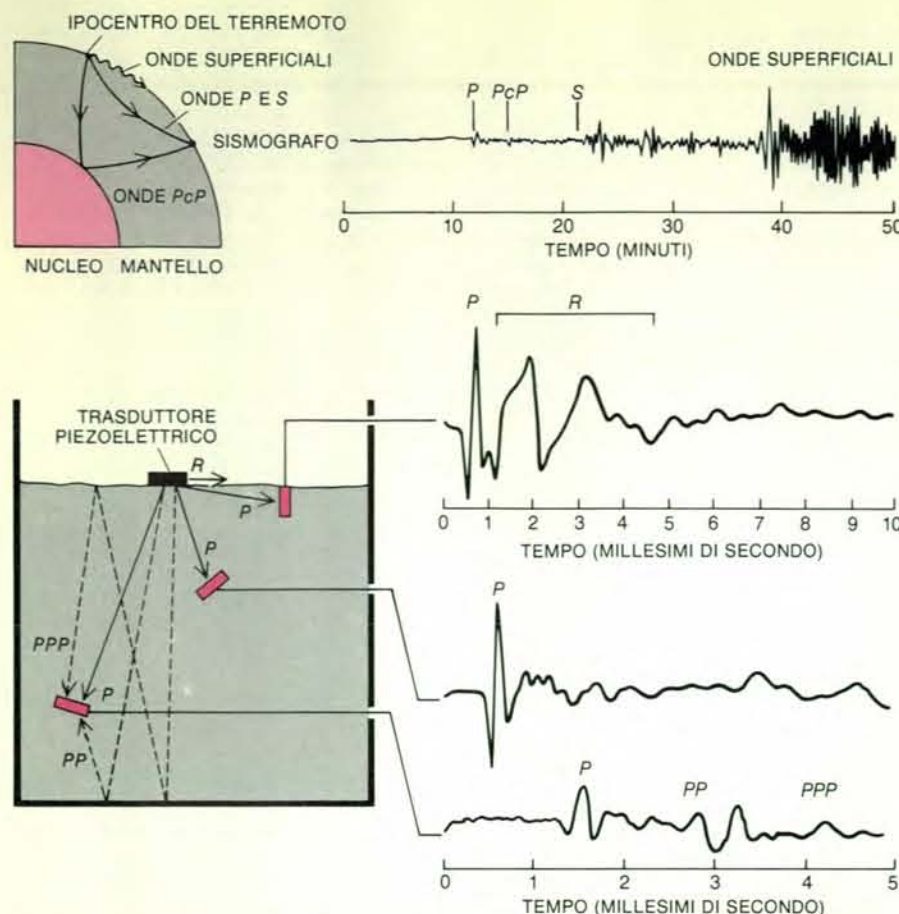
Una serie di semplici esperimenti che sfruttavano questa risposta di orientamento difensivo hanno confermato la nostra impressione secondo la quale non erano gli stimoli visivi e acustici a guidare tale comportamento. Il fatto di ricoprire tutti gli otto occhi dell'animale con colorante opaco non ha avuto il minimo effetto, né sulla sua sensibilità agli stimoli di minaccia né sulla precisione con cui l'animale si girava nella loro direzione. Nemmeno la frapposizione di pannelli fonoassorbenti fra stimolo e scorpione ha influenzato la risposta di quest'ultimo.

Per contro, quando ho bloccato la trasmissione delle onde meccaniche attraverso la sabbia inserendo un cuscino d'aria, il comportamento dell'animale si è modificato, anche se il dispositivo non interferiva nei segnali visivi e acustici. Quando uno scorpione eccitato veniva posto a un lato del cuscino d'aria e all'altro lato si emetteva uno stimolo, l'animale non reagiva nemmeno se l'origine della perturbazione era distante pochi centimetri. Ma se almeno una delle sue zampe toccava il substrato sul lato perturbato, l'animale si voltava immediatamente in quella direzione. Perché lo scorpione individui uno stimolo, deve essere, dunque in contatto meccanico con il substrato.

Inoltre l'angolo della risposta difensiva dipendeva unicamente dalla combinazione di zampe che entrava in contatto con la superficie stimolata. Per esempio, quando tutte le zampe di sinistra erano da un lato del cuscino d'aria e tutte le zampe di destra erano dall'altro lato, la stimolazione della sabbia da varie angolazioni a sinistra dello scorpione provocava sempre risposte a sinistra, ma con una forte tendenza per la direzione perpendicolare al cuscino d'aria. Quest'ultimo non diminuiva la sensibilità dello scorpione alle vibrazioni, ma riduceva la precisione con cui l'animale riusciva a localizzare la direzione della sorgente delle vibrazioni. Queste risposte indicavano che, di norma, lo scorpione determina l'angolo di cui si gira integrando le informazioni provenienti da tutte le zampe. Eliminando selettivamente una parte di queste informazioni, avevo sistematicamente scoordinato la percezione della direzione della sorgente.

Gli esperimenti con il cuscino d'aria hanno dimostrato che lo scorpione acquisisce le informazioni necessarie a localizzare la preda dai segnali trasmessigli attraverso il substrato e che probabilmente identifica questi segnali grazie a recettori situati nelle zampe. A questo punto rimaneva da determinare ancora la natura di questi recettori e, prima di tutto, la natura dei segnali meccanici che, condotti dalla sabbia, portano l'informazione.

Quest'ultimo problema era particolarmente interessante poiché i geofisici hanno sempre considerato la sabbia un cattivo conduttore di onde meccaniche per due ragioni. Innanzitutto la sabbia è un mezzo granulare disaggregato, in cui le particelle contigue non sono unite insieme in maniera rigida. In un aggregato solido come una roccia, una sollecitazione applicata a una particella si trasferisce a quella successiva e lo spostamento si propaga come un'onda elastica, mentre nella sabbia l'energia cinetica si esaurisce rapidamente per attrito, dato che i granelli possono scorrere reciprocamente. Questa forma di anelasticità tende a smorzare, in particolare, le oscillazioni di bassa frequenza e di grande lunghezza d'onda.

D'altro canto, alle frequenze più elevate, le lunghezze d'onda diventano talmente corte da avvicinarsi alle dimensioni dei granelli di sabbia; questi riflettono e dif-

Gli esperimenti con il cuscino d'aria dimostrano che, per localizzare la preda, gli scorpioni reagiscono alle vibrazioni del substrato più che a stimoli visivi o acustici. Uno scorpione posto a un lato del cuscino d'aria, che blocca la trasmissione delle vibrazioni attraverso la sabbia, non manifesta risposta a una perturbazione distante parecchi centimetri sul lato opposto *(a)*. Quando le sue zampe di sinistra sono in contatto con il substrato stimolato *(b)*, lo scorpione si volge verso sinistra, ma l'angolo di risposta propende fortemente a coincidere con la direzione a 90 gradi (perpendicolare al cuscino d'aria), indipendentemente dall'angolo della sorgente della perturbazione *(c)*. Questo indica che, per valutare con precisione la direzione della preda, è indispensabile che tutte le otto zampe trasmettano uno stimolo sensoriale.

Le onde meccaniche provocate da un terremoto sono rilevabili da un sismografo a migliaia di chilometri dall'ipocentro *(in alto)*; gli stessi tipi di onda si propagano nella sabbia su distanze brevi in risposta alle tenui vibrazioni di un trasduttore piezoelettrico *(in basso)*. I terremoti emettono due tipi di onde interne, longitudinali o di compressione *(P)* e trasversali *(S)*, e due tipi di onde superficiali, le onde di Rayleigh e le onde di Love. Le onde *P* sono le più veloci e quindi arrivano prima ai sismografi, seguite dalle loro riflessioni dal nucleo *(PcP)* e poi dalle onde trasversali e superficiali. La sabbia conduce unicamente onde di compressione e di Rayleigh, che vengono entrambe registrate da un sensore piezoelettrico *(in colore)* posto alla superficie della cassetta piena di sabbia. A profondità di sei centimetri e di 19,5 centimetri vengono rilevate solo le onde *P*, ma la sabbia le conduce abbastanza bene da poterne misurare le riflessioni.

fondono quei segnali, filtrando efficacemente dall'onda che si propaga le componenti ad alta frequenza. Per questi due effetti (attenuazione della bassa frequenza e propagazione dell'alta frequenza), la sabbia trasmetterà le onde meccaniche solo entro un intervallo ben limitato di frequenze, ammesso che lo faccia.

Eppure era evidente che gli scorpioni ricevessero dei segnali. Ho così deciso di analizzare in maggior dettaglio le proprietà di conduzione della sabbia. A questo scopo ho disposto a varie profondità e sulla superficie di una cassetta di sabbia alcuni sismografi piezoelettrici. Ho poi utilizzato un altro trasduttore piezoelettrico per generare un impulso intermittente nel substrato, a una certa distanza dai sismografi. Le prime registrazioni effettuate hanno confermato quanto il comportamento dello scorpione ci aveva suggerito: su distanze di anche parecchi decimetri, la sabbia è un conduttore di vibrazioni meccaniche abbastanza buono.

Rimaneva da determinare però che cosa esattamente stessero registrando questi sismografi. I tipi di onde elastiche che si possono propagare nei solidi sono quattro. Due di essi, le onde longitudinali (o di compressione) e le onde trasversali, sono onde interne e si propagano come fronti che si espandono sfericamente per tutto il mezzo. Le onde di compressione (onde sonore) fanno oscillare le particelle avanti e indietro lungo la direzione di propagazione dell'onda, mentre le onde trasversali fanno muovere le particelle perpendicolarmente alla direzione di propagazione. Gli altri due tipi di perturbazione elastica si propagano in superficie. Di queste, sono interessanti ai nostri fini solo le onde di Rayleigh, che fanno muovere le particelle lungo un'ellisse retrograda, il cui piano è perpendicolare alla superficie e parallelo alla direzione di propagazione.

I rilevatori piezoelettrici hanno registrato due forti segnali propagantisi attraverso la cassetta di sabbia: un'onda relativamente veloce sia nella sabbia sia in superficie e un'onda più lenta registrata solo in superficie. Orientando l'asse sensibile dei rilevatori ad angoli diversi rispetto alla direzione di propagazione del segnale, ho potuto identificare i movimenti delle particelle associati a ciascuna delle due onde. L'onda veloce è risultata un'onda di compressione, mentre quella lenta, superficiale, era un'onda di Rayleigh. Non sono stati rilevati movimenti di particelle indicanti propagazione di onde trasversali, il che non sorprende: la conduzione di questo tipo di onde dipende interamente dalle forze restauratrici della forma, fra particelle contigue, ed è probabile che queste forze siano minime in un mezzo granulare come la sabbia.

Contrariamente a quanto si crede, la sabbia non è quindi un mezzo con notevole capacità di assorbimento, ma conduce almeno due tipi di onda. Inoltre, le proprietà fisiche di questi segnali - intervallo di frequenze e velocità di propagazione - erano tali che animali dotati di recettori con sensibilità appropriata avrebbero potuto utilizzarle per rilevare e raccogliere informazioni a distanze di parecchi decimetri.

Per determinare la facilità di propagazione nella sabbia di onde di frequenza e lunghezza d'onda diverse ho analizzato le forme d'onda di segnali di compressione registrati a distanze differenti da una emittente costante. Questo procedimento, che è in pratica una analisi di Fourier, consente di determinare il grado di attenuazione delle varie frequenze che compongono un segnale complesso che si propaga attraverso il mezzo. Ho trovato che la sabbia è un conduttore piuttosto buono di onde meccaniche nella banda di frequenza tra 0,1 e 5 chilohertz. In particolare, le frequenze prevalenti dei segnali registrati relativamente lontano dalla fonte si approssimavano a un chilohertz, una frequenza alla quale è molto sensibile la maggioranza dei meccanorecettori animali. Per di più, nella sabbia, le lunghezze d'onda corrispondenti a queste frequenze sono sufficientemente corte (alcuni centimetri) da consentire a piccoli animali come gli scorpioni di distinguere tempi diversi di arrivo dell'onda a recettori differenti.

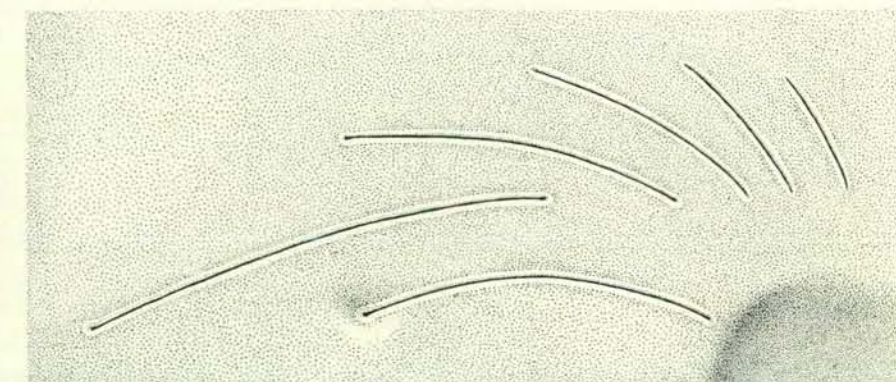
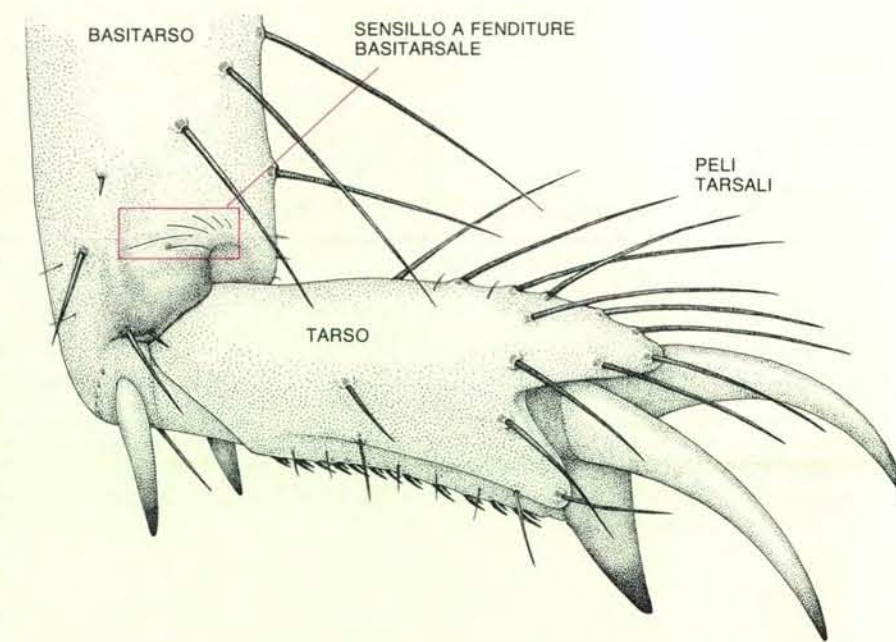
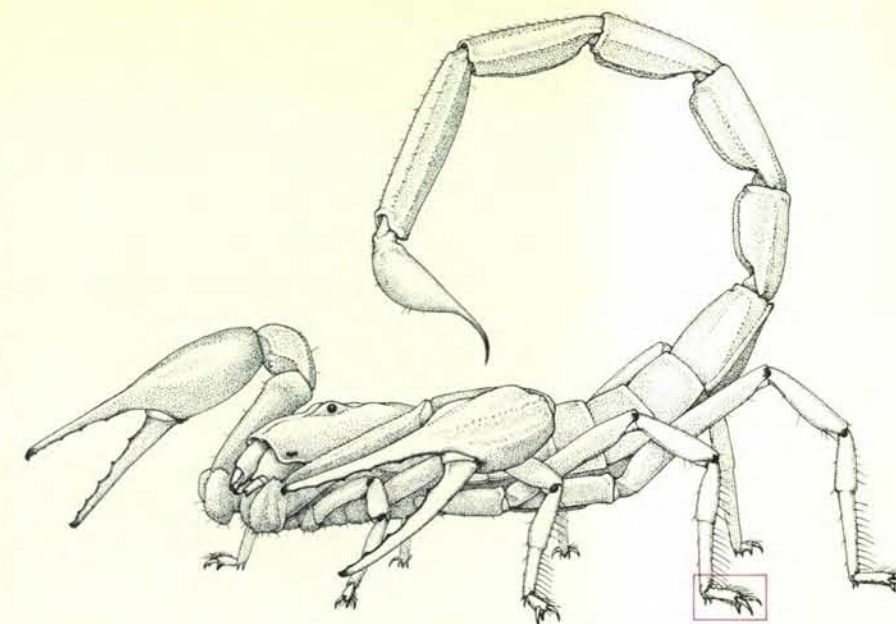
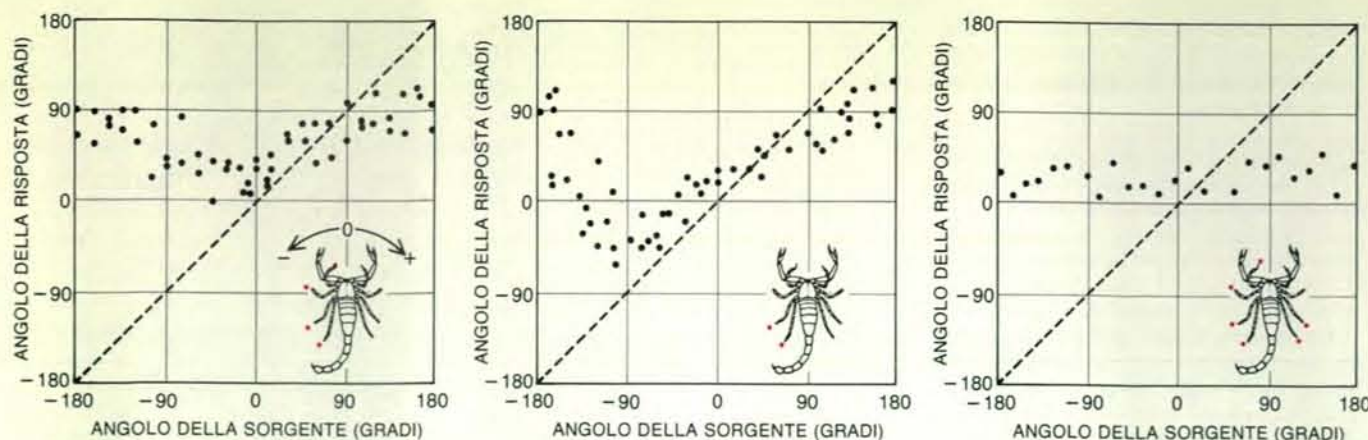
La caratteristica più sorprendente dei segnali era forse la bassa velocità di trasmissione. La velocità elevata delle onde elastiche nella maggioranza degli altri solidi naturali è una delle ragioni principali per cui gli studiosi di biologia del comportamento sono stati sempre propensi a trascurare le vibrazioni del substrato come potenziali stimoli di orientamento. L'alta velocità, così come la grande lunghezza d'onda, impedisce a un animale di rilevare le differenze nel tempo di arrivo o nell'ampiezza dei segnali a livello dei diversi recettori e, di conseguenza, di utilizzare queste caratteristiche per localizzare la sorgente. Le misurazioni hanno però dimostrato che la velocità delle onde di compressione nella sabbia superficiale non compatta è solo di circa 120 metri al secondo, cioè meno della metà della velocità del suono nell'aria; nella sabbia compatta, invece, è di circa 200 metri al secondo, ossia approssimativamente di un ordine di grandezza inferiore alla velocità dell'onda sonora nella maggior parte degli altri solidi. Nella sabbia, le onde di Rayleigh si propagano ancora più lentamente, addirittura a 40-50 metri al secondo.

Sia le onde di compressione sia le onde di Rayleigh potrebbero quindi essere rilevate dagli scorpioni e sfruttate per trovare la preda. La fase successiva della nostra ricerca è consistita nel trovare quali vibrazioni questi animali avvertivano effettivamente e con quali recettori.

I segmenti tarsali della zampa dello scorpione erano il punto più opportuno in cui cercare i recettori per le vibrazioni, dato che i nostri studi avevano dimostrato che essi sono l'unica parte del corpo che deve essere in contatto con il terreno perché l'animale possa orientarsi. Ciascuna delle otto zampe porta parecchie strutture cuticolari che potrebbero funzionare da meccanorecettori. Queste strutture includono sensilli in forma di aculei e uncini e una struttura particolare comune a tutti gli aracnidi, un sensillo costituito da fenditure, l'organo più probabile per essere un meccanorecettore, poiché è noto che la struttura omologa presente nei ragni, l'organo liriforme, è responsabile della straordinaria sensibilità di questi aracnidi alle vibrazioni della tela.

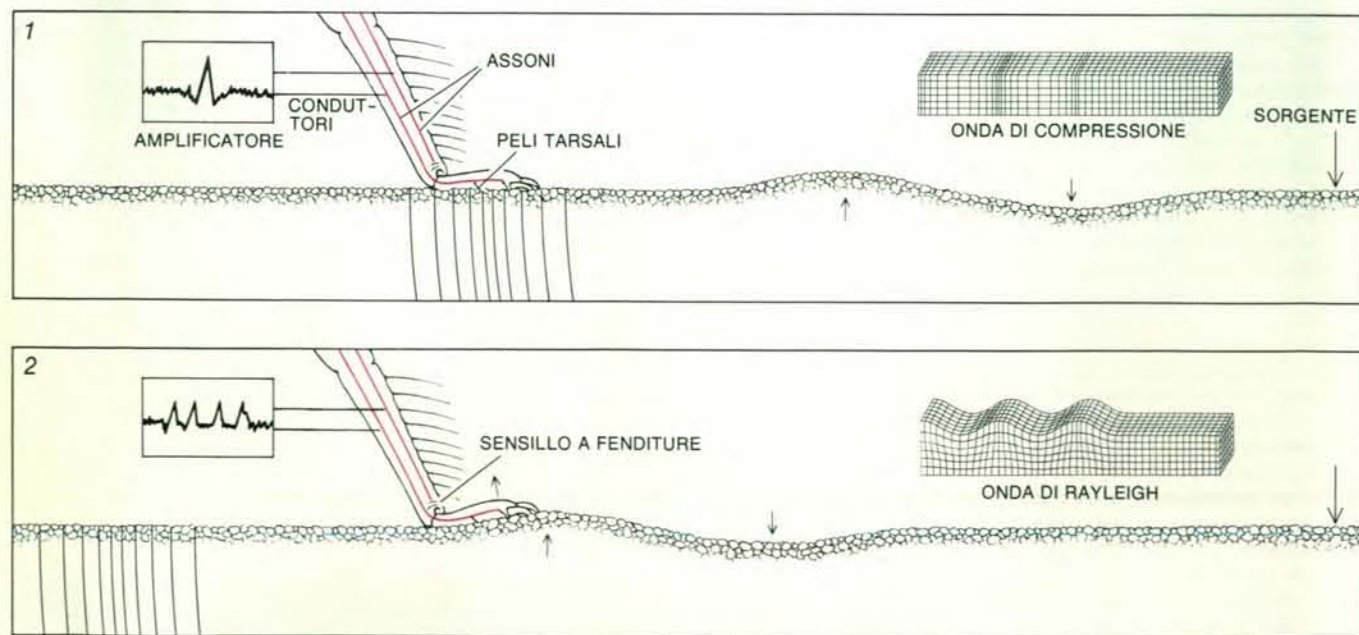
Non è stato difficile dimostrare che i sensilli a fenditure, presenti nei tarsi, sono di fatto una componente di cruciale importanza nel meccanismo che consente allo scorpione di localizzare la preda. Quando abbiamo perforato con un ago sottile tutti i sensilli a fenditure di uno scorpione, l'animale si è comportato in modo del tutto normale, tranne per il fatto che diventava completamente insensibile a tutti gli stimoli eccettuate le perturbazioni più intense del substrato. D'altro canto, quando abbiamo perforato solo i sensilli delle zampe di sinistra, lo scorpione si è comportato in modo molto simile a quando si trovava «a cavalcioni» del cuscino d'aria, ossia a ciascuna sollecitazione del substrato tendeva a girarsi di 90 gradi verso destra, cioè dal lato delle zampe intatte. La somiglianza di questi due schemi di comportamento indica che la distruzione dei sensilli a fenditure sul lato sinistro del corpo dello scorpione corrisponde alla eliminazione di qualsiasi stimolo sensoriale che interessi quel lato.

Analogamente, quando abbiamo reciso soltanto due fenditure adiacenti, la risposta di orientamento è rimasta sufficientemente precisa, salvo nei casi in cui i bersagli fossero proprio vicino a quelle fenditure. Quando, infine, abbiamo eliminato tutti i sensilli eccetto due adiacenti, lo scorpione ha continuato a rivolgersi nella direzione dei sensilli rimasti, indipendentemente dalla posizione del bersaglio. Era chiaro, dunque, che per poter determinare con precisione la direzione della fonte della perturbazione lo scorpione aveva bisogno di uno stimolo sensoriale proveniente da tutti i sensilli a fenditure.

Dato che le zampe dello scorpione sono relativamente grosse, è stato possibile inserire conduttori estremamente sottili nei segmenti terminali e registrare così i

Due tipi di meccanorecettori sul segmento tarsale (terminale) della zampa dello scorpione sono sensibili alle tenui vibrazioni del substrato. I peli che si protendono dai lati e dalla base del tarso poggiano sopra i granelli di sabbia e si infilano tra di essi. Il sensillo basitarsale a fenditure presenta zone nelle quali la cuticola si ripiega su se stessa; esso è particolarmente sensibile alle vibrazioni che comprimono le fenditure in direzione perpendicolare al loro asse maggiore ed è in grado di rilevare movimenti del substrato che abbiano un'ampiezza di circa 0,1 nanometri (10^{-10} metri).

La funzione fondamentale dei sensilli a fenditure nella determinazione della direzione della sorgente delle vibrazioni è indicata da esperimenti in cui alcuni di questi sensilli o tutti vengono perforati *(punti in colore)* e il substrato è poi stimolato ad angolazioni diverse rispetto allo scorpione. Quando si asportano i sensilli su tutte le otto zampe, lo scorpione è «accecato», salvo per quanto riguarda le vibrazioni più intense del substrato. L'ablazione selettiva induce, nella reazione dello scorpione, una propensione a girarsi (dalla parte delle zampe intatte) con un angolo che aumenta in proporzione al numero dei sensilli asportati. Per esempio, quando sono danneggiate le quattro zampe di sinistra *(a sinistra)* lo scorpione si volta sempre verso destra, anche quando la sollecitazione proviene da sinistra. È probabile quindi che lo scorpione normalmente integri le informazioni provenienti dai sensilli a fenditure di tutte le otto zampe per determinare la direzione della preda.

potenziali d'azione, o segnali bioelettrici, trasmessi dai recettori al cervello. Queste registrazioni hanno messo in evidenza potenziali d'azione di due dimensioni: quelli maggiori comparivano durante la fase iniziale della risposta alla stimolazione vibratoria ed erano associati all'arrivo nel tarso delle onde di compressione, relativamente veloci. I potenziali d'azione minori si manifestavano in fasi successive, in concomitanza con l'arrivo delle più lente onde di Rayleigh.

Stimolando selettivamente i singoli meccanorecettori tarsali ho potuto quindi isolare quelli responsabili dei due distinti segnali bioelettrici. Questi esperimenti hanno messo in evidenza che il potenziale di maggiore intensità è una risposta alla stimolazione dei peli sensoriali situati sul tarso, mentre quello più debole ha origine nei sensilli a fenditure. Ne consegue che i peli sensoriali rilevano le onde di compressione e i sensilli a fenditure l'arrivo delle onde di Rayleigh.

La differenza di sensibilità fra i due tipi di organi sensoriali è comprensibile se si tiene conto dei diversi movimenti di particelle associati alle onde che tali organi registrano. Friedrich G. Barth e collaboratori, dell'Università Johann Wolfgang Goethe di Francoforte, hanno dimostrato che i sensilli a fenditure sono particolarmente sensibili alle forze che comprimono le fenditure in direzione perpendicolare al loro asse maggiore. È proprio quello che farebbe un'onda di Rayleigh, dato che il suo moto verticale sposta verso l'alto il segmento terminale della zampa. Per contro, i peli sensoriali si protendono dai lati e dalla superficie inferiore della zampa nella sabbia e quindi sono nella posizione adatta per rilevare il moto orizzontale delle particelle provocato dalle onde di compressione.

La risposta meccanosensoriale di uno scorpione alle onde che si propagano attraverso la sabbia può essere osservata inserendo sottili conduttori metallici nei segmenti terminali delle zampe e registrando i segnali bioelettrici (potenziali d'azione) trasmessi lungo il nervo della zampa. Un sensore piezoelettrico (non riprodotto) posto nella sabbia in prossimità del tarso registra le onde. Quando la sabbia viene smossa, i primi segnali che arrivano al tarso sono le onde di compressione *(1)*, che stimolano i peli tarsali provocando l'ascesa lungo il nervo della zampa di potenziali d'azione molto ampi. Dopo qualche millesimo di secondo *(2)*, il moto verticale delle particelle associato con le onde di Rayleigh più lente comprime il sensillo a fenditure, provocando segnali di minore ampiezza. Sembra che la stimolazione dei sensilli a fenditure da parte delle onde di Rayleigh sia alla base della percezione della direzione del bersaglio. L'animale può valutare la distanza confrontando l'intervallo di tempo fra successive stimolazioni dei due tipi di meccanorecettori da parte dei due tipi di onda.

Localizzare la preda non consiste solo, però, nel semplice rilevamento del movimento della sabbia; lo scorpione riesce infatti, in qualche modo, a ricavare informazioni dalle onde e a utilizzarle per determinare la posizione della sorgente della perturbazione. Dato che la funzione dei sensilli a fenditure è essenziale per determinare la direzione di quella sorgente, è probabile che le vibrazioni delle onde di Rayleigh costituiscano l'indicazione giusta che consente di determinare l'angolazione del bersaglio. Finora né l'anatomia né la risposta fisiologica sia dei sensilli a fenditure sia dei peli sensoriali ha potuto essere considerata influente ai fini dell'orientamento. In altri termini, i recettori che, nello scorpione, captano le vibrazioni reagiscono in ugual misura a onde provenienti da qualsiasi angolazione. Quindi la determinazione della direzione della preda da parte dell'animale deve comportare un confronto fra le risposte di due recettori separati spazialmente, quando un'onda passa sotto l'animale.

Le otto zampe dello scorpione adulto formano un «campo sensoriale» approssimativamente circolare, con un diametro fra quattro e sei centimetri circa. Lo scorpione potrebbe sfruttare questo spazio relativamente ampio in uno o l'altro di due modi. In primo luogo, potrebbe determinare la direzione di una sorgente d'onda dall'intervallo di tempo che intercorre fra la stimolazione dei recettori più vicini alla sorgente e dei recettori più lontani; lo scorpione potrebbe semplicemente volgersi nella direzione dei recettori che sono stati stimolati per primi. Se si suppone che il campo sensoriale abbia un diametro di cinque centimetri, l'intervallo sarebbe di circa un millesimo di secondo per un'onda di Rayleigh che si propaghi alla velocità di 50 metri al secondo, mentre per un'onda di compressione con velocità di 150 metri al secondo esso sarebbe di circa 0,3 millesimi di secondo. Molti animali utilizzano intervalli di tempo inferiori per localizzare la sorgente di onde di compressione che si propagano nell'aria; gli essere umani, per esempio, possono facilmente stabilire la direzione di una sorgente sonora sulla base di un intervallo di tempo fra la stimolazione dei due orecchi inferiore a 10 microsecondi.

Lo scorpione potrebbe anche stabilire la direzione di una sorgente d'onda in base alla differenza d'intensità con cui l'onda stimola recettori differenti. Quando un'onda si propaga la sua ampiezza diminuisce, in parte perché il fronte d'onda si espande geometricamente diffondendo l'energia dell'onda e in parte perché il segnale è assorbito dal mezzo. I recettori più vicini alla sorgente dovrebbero quindi essere stimolati più intensamente.

Per verificare quale delle due indica-

Lo scorpione percepisce la direzione della fonte della perturbazione, determinando quali delle sue zampe sono stimolate per prime da un'onda di Rayleigh. Nell'esperimento illustrato lo scorpione è disposto a cavallo di una fessura che divide una piattaforma, in modo che le zampe di sinistra possano essere stimolate indipendentemente dalle zampe di destra. Un oscilloscopio registra gli impulsi applicati a ciascun lato della piattaforma. Lo scorpione si rivolge sempre nella direzione delle zampe che sono state stimolate per prime, anche se le altre ricevono sollecitazioni più intense. Questo dimostra che è l'intervallo di tempo fra la stimolazione delle zampe opposte e non l'intensità relativa della stimolazione a codificare l'informazione utilizzata per determinare la direzione della sorgente. Questi esperimenti con la piattaforma dimostrano anche che gli scorpioni possono rilevare intervalli di tempo di anche solo 0,2 millesimi di secondo; rispondono tuttavia in maniera più coerente a intervalli da uno a due millesimi di secondo, pressappoco il tempo occorrente a un'onda di Rayleigh per attraversare la distanza fra le due serie di zampe.

zioni sia più importante per lo scorpione, ho costruito una piattaforma divisa longitudinalmente nel mezzo da una stretta fessura; ciascuna metà della piattaforma poteva vibrare in modo autonomo mediante propri stimolatori meccanici. Disponendo uno scorpione a cavallo della fessura, il tempo e l'intensità delle vibrazioni che raggiungevano le sue zampe di sinistra e di destra potevano essere fatte variare indipendentemente.

Questo accorgimento mi ha consentito di porre lo scorpione dinnanzi a un dilemma. In condizioni naturali, le zampe più vicine a una sorgente d'onda sono stimolate per prime e più intensamente. Invece, con la piattaforma divisa, è stato possibile eccitare per prime le zampe di destra, ma con una intensità pari alla metà di quella della stimolazione inferta alle zampe di sinistra, costringendo così lo scorpione a scegliere lo stimolo più importante. In ripetute prove, lo scorpione si voltava sempre verso destra, indipendentemente dall'intensità dello stimolo che arrivava alle zampe di sinistra. Si è visto così che l'indicazione utilizzata dall'animale per determinare la direzione della sorgente non è l'intensità relativa, ma il tempo di arrivo.

Per di più, lo scorpione sapeva riconoscere benissimo quali zampe venivano stimolate per prime; intervalli di tempo compresi fra uno e due millesimi di secondo riuscivano a simulare al meglio un'onda naturale e questo era provato dalla velocità e dalla precisione con cui lo scorpione si voltava verso destra o verso sinistra; un millesimo di secondo è pressappoco il tempo che impiega un'onda superficiale per percorrere i cinque centimetri di diametro del campo sensoriale di uno scorpione. Ma anche intervalli di soli 0,2 millesimi di secondo provocavano risposte precise.

Rimane da spiegare in che modo lo scorpione percepisce la distanza dalla preda. Le nostre prime osservazioni sul campo avevano dimostrato che l'animale può valutare la distanza fino a 15 centimetri e che raramente sbaglia a 10 centimetri o a distanze inferiori. Purtroppo, dato che in laboratorio lo scorpione è restio a cacciare, non è facile studiare il meccanismo che utilizza per percepire la distanza, cosicché rimane ancora confinato nel campo delle ipotesi. Tuttavia, come nel caso della determinazione della direzione della sorgente, la determinazione della distanza deve comportare il rilevamento di piccole differenze nel tempo o nell'ampiezza della stimolazione di recettori diversi.

È anche probabile che l'informazione fornita dalle onde di compressione, che non sembra contribuire alla risposta di orientamento, sia un fattore importante nella percezione della distanza. L'animale potrebbe stabilire l'intervallo di tempo fra la stimolazione dei peli tarsali da parte delle onde di compressione e la stimolazione del sensillo a fenditure indotta dalle onde di Rayleigh, più lente. Tale intervallo sarebbe proporzionale alla distanza della sorgente. Esso potrebbe anche avvertire l'attenuazione di questi segnali; l'ampiezza delle onde meccaniche diminuisce rapidamente a mano a mano che queste procedono attraverso la sabbia e ciò vale in particolar modo per le onde di compressione interne, poiché la loro energia si dissipa lungo un fronte d'onda sferico, e non su un fronte circolare, caratteristico delle onde superficiali. Quando uno scorpione è vicino a una sorgente d'onda, dovrebbe avvertire una differenza maggiore nell'intensità dello stimolo fra zampe vicine e lontane, di quando ne è lontano.

In retrospettiva, sembra contraddire il senso comune il fatto che un animale, particolarmente se vive nella sabbia, debba dipendere in modo così esclusivo da informazioni trasmesse attraverso il terreno. Tuttavia l'utilizzazione da parte di *Paruroctonus mesaensis* di onde che si propagano in un solido non è probabilmente un caso unico. Può darsi che altri substrati naturali - una foglia, una mela, un tronco d'albero - trasmettano, in virtù della loro forma e delle loro caratteristiche fisiche particolari, informazioni importanti su brevi distanze, adeguate al mondo ristretto degli animali che vivono in essi. In effetti, *Paruroctonus* è solo un esempio preso da un elenco sempre più lungo di organismi, che si orientano verso la sorgente del movimento in un substrato solido utilizzando le zampe per «ascoltare» oltre che per camminare.

Il pozzo più profondo del mondo

La perforazione del pozzo di Kola, giunta ora a 12 000 metri, ha rivelato la causa di una discontinuità sismica e ha contribuito allo sviluppo di nuove tecniche per l'esplorazione profonda della crosta della Terra

di Ye. A. Kozlovsky

Dal 1970, nelle nude rocce della Penisola di Kola, 250 chilometri a nord del circolo polare artico, nella regione sovietica di Murmansk, una torre di trivellazione alta come un edificio di 27 piani, sta perforando un pozzo nello scudo continentale baltico. La punta della trivella (o scalpello) che è avanzata per più di metà percorso in rocce cristalline, è arrivata ora a una profondità di oltre 12 000 metri, cosicché il pozzo di Kola è il più profondo del mondo. Supera di gran lunga i pozzi, profondi da 1500 a 7000 metri, utilizzati di solito per l'esplorazione o per l'estrazione di petrolio, carbone, ferro, metalli non ferrosi, diamanti o altre preziose risorse della Terra.

Il primo obiettivo importante del pozzo di Kola è la comprensione della struttura profonda della crosta continentale e delle forze che l'hanno modellata nel corso di quattro miliardi di anni di storia geologica. Le conoscenze attuali derivano in gran parte dallo studio di rocce che affiorano qua e là e che nel complesso riguardano circa 15 000 metri dello spessore totale della crosta che è in media di 30 000 metri. Molte altre informazioni sulla natura della crosta nascosta sono fornite dalla velocità delle onde sismiche, che aumenta con la profondità, ma che varia con la composizione della roccia attraverso cui le onde si propagano, e dai dati raccolti da strumenti a bordo di aerei e di satelliti, che rilevano il campo gravitazionale ed elettromagnetico della Terra. Tuttavia nulla può sostituire l'osservazione diretta in profondità delle strutture e dei processi che vi hanno luogo.

Il pozzo di Kola ha attraversato 1,4 miliardi di anni di storia della Terra, passando da rocce del Proterozoico a rocce dell'Archeano di età compresa tra 2,5 e 2,7 miliardi di anni. Esso ha messo allo scoperto diversi cicli di formazione della crosta attraverso i quali nuova roccia ignea (ossia formata per consolidamento di magma) è stata portata nella crosta dal mantello fuso sottostante, poi frantumata dall'azione degli agenti atmosferici e dalle glaciazioni, quindi divisa e ridistribuita in strati sedimentari e, infine, riconsolidata in roccia cristallina per metamorfismo, dovuto al calore e alla pressione derivanti da successive intrusioni di roccia ignea. I campioni di roccia, portati in superficie con sempre maggiore difficoltà all'aumentare della profondità del pozzo, hanno fornito un modello termico generale dell'evoluzione della crosta nell'intervallo di tempo in cui si sono delineate le principali configurazioni degli scudi continentali della Terra. La documentazione raccolta permette di stabilire, per lo meno, che la crosta primitiva della Penisola di Kola è granodiorite, una roccia intrusiva con un contenuto in quarzo un poco inferiore al granito.

Uno dei principali obiettivi del pozzo di Kola era quello di attraversare lo strato di crosta superiore di granito e di arrivare al sottostante basamento roccioso di composizione basaltica. I basalti non si trovano frequentemente in superficie e il presunto basamento basaltico, poi, non si trova mai sui continenti. Si ritiene che il passaggio tra i due tipi di crosta sia indicato da un brusco aumento della velocità delle onde sismiche, osservato in tutto il mondo a profondità intermedie della crosta; questa variazione di velocità, nella Penisola di Kola, avviene a 9000 metri di profondità. Tuttavia il pozzo di Kola, che è stato il primo a superare tale limite, al di sotto non ha incontrato basalto.

Il pozzo ha invece raggiunto la base di una zona anomala di rocce metamorfiche disgregate che aveva incontrato per la prima volta a 4500 metri. Penetrando in questa zona il pozzo ha incontrato flussi di acqua calda fortemente mineralizzata, sorprendentemente abbondanti. Quest'acqua (detta «acqua di cristallizzazione») deriva dai minerali costituenti le rocce cristalline, viene liberata quando i componenti delle rocce sono soggetti a fusione e a ricristallizzazione durante i processi metamorfici e svolge un ruolo fondamentale nella genesi dei giacimenti minerari. Di solito quest'acqua fluisce via dalla formazione metamorfica e va a depositare il suo carico di minerali in una zona più superficiale della crosta. A Kola l'acqua è stata confinata nello stesso punto in cui è stata liberata da strati sovrastanti impermeabili di rocce ignee spessi 2000 metri. Per risospingere l'acqua nella roccia sarebbe necessaria una pressione molto elevata che si ha solo a maggiori profondità nella crosta o nel mantello superiore. Poiché la resistenza della roccia alla frattura è solo una frazione di questa pressione idraulica, la disidratazione della roccia è stata accompagnata da una microfratturazione. Questo fenomeno di disgregazione idraulica della roccia metamorfica, mai osservato in precedenza, può avere una funzione significativa nella struttura dei continenti.

I flussi di acqua mineralizzata rinvenuti nella zona di disgregazione, le sottili e ben definite formazioni sottostanti e la scoperta di depositi di minerali ad altri livelli del pozzo di Kola dimostrano che l'uomo, nella ricerca mineraria, ha appena scalfito la superficie terrestre: immense risorse giacciono in profondità in attesa di essere raggiunte dalla tecnologia umana.

La perforazione del pozzo ha inoltre evidenziato flussi di gas a tutti i livelli. Tra i gas identificati vi sono elio, idrogeno, azoto, metano e altri idrocarburi, infine anidride carbonica. Gli elementi leggeri presenti in essi sono derivati dalla struttura cristallina della roccia in seguito agli stessi processi metamorfici che hanno liberato l'acqua. La composizione isotopica del carbonio suggerisce tuttavia che le fonti di anidride carbonica siano state due, di cui una biogena connessa alla presenza di microrganismi fossili nelle rocce del Proterozoico che risalgono a centinaia di milioni di anni fa.

La perforazione del pozzo di Kola ha inaugurato, in territorio sovietico, un programma a lungo termine di sondaggi sistematici della struttura profonda della crosta terrestre, che proseguirà nel prossimo secolo. Nel 1962, sotto gli auspici del Consiglio interdipartimentale per lo studio dell'interno della Terra e per le perforazioni a grande profondità, geologi e ingegneri minerari hanno cominciato a sviluppare la tecnologia per la perforazione e lo studio dei pozzi profondi. Nel 1970 hanno avuto

Vicino alla testa del pozzo di Kola sono pronte le aste di perforazione in lega di alluminio molto resistente, che verranno a mano a mano utilizzate. La fotografia è stata ripresa all'interno della torre di trivellazione, una struttura alta circa 86 metri, che si innalza al di sopra della testa del pozzo e che sostiene la batteria di perforazione, con una capacità di carico di 400 tonnellate. Il gruppo motore che è situato al fondo del pozzo è costituito da una turbina che viene azionata dal flusso ad alta pressione del fango di perforazione pompato dalla superficie.

La costruzione che ospita la torre di trivellazione si innalza come un edificio di 30 piani nella tundra della Penisola di Kola, 250 chilometri a nord del circolo polare artico. La copertura protegge la testa del pozzo dal gelo, permettendo il funzionamento dell'impianto per tutto l'anno.

I siti scelti per la perforazione, destinata a uno studio di vasta portata della crosta terrestre, punteggiano il territorio dell'Unione Sovietica. Grazie alla tecnologia messa a punto nella perforazione del pozzo di Kola, profondo 12 000 metri, e del pozzo di Saatly, profondo 8500 metri, sul Mar Caspio, verranno perforati tre nuovi pozzi profondi più di 7000 metri nelle località indicate nella cartina. Contemporaneamente verranno perforati pozzi meno profondi nei siti indicati dai pallini in colore privi di nome. Tutti i pozzi si trovano nei punti d'intersezione di profili sismici recenti e serviranno a comprendere la genesi dei depositi di minerali e di idrocarburi.

inizio le perforazioni a Kola e a Saatly, nel distretto petrolifero di Baku sul Mar Caspio, dove il pozzo ha raggiunto gli 8500 metri. Nel frattempo una prospezione sismica profonda del territorio sovietico ha contribuito alla scelta dei siti per altri pozzi profondi e superprofondi.

Una migliore conoscenza della struttura profonda della crosta servirà a trovare i metodi per individuare ed esplorare depositi di minerali e campi di idrocarburi a profondità molto maggiori di quelle raggiunte finora. Infatti, come ha dimostrato il pozzo di Kola, risorse minerarie possono trovarsi anche a grandi profondità. Un obiettivo fondamentale di questa impresa di ricerca è quindi quello di sviluppare la tecnologia che permetta di penetrare nell'interno della Terra fino a 10000-15000 metri.

Sulla scala globale la crosta terrestre è piuttosto sottile: solo 35 chilometri su un raggio di 6000 chilometri. Sembra che i dati sismici e altri dati di tipo indiretto confermino l'ipotesi, avanzata nel 1926 dal geofisico inglese Harold Jeffreys, secondo la quale la crosta continentale sarebbe divisa in tre strati: sedimentario, granitico e basaltico. La correlazione tra densità dei campioni di roccia misurata in laboratorio e velocità delle onde sismiche osservata nelle prospezioni, dà per le rocce sedimentarie, una densità di 1,8-2,5 grammi per centimetro cubo e una velocità di circa cinque chilometri al secondo. Densità e velocità aumentano rispettivamente a un valore tra 2,5 e 2,75 grammi per centimetro cubo e tra cinque e sei chilometri al secondo nelle rocce granitiche. In corrispondenza della discontinuità di Conrad la velocità delle onde sismiche raggiunge i sei - sette chilometri al secondo; questo valore è stato preso per indicare la transizione allo strato basaltico in cui la densità è di 2,75-3 grammi per centimetro cubo. Lo strato granitico, costituito da elementi più leggeri, come ossigeno, silicio e alluminio, fu messo in posto nell'Archeano ed è ampiamente sviluppato sulla superficie dei continenti. Mentre la crosta continentale può raggiungere uno spessore di 70-75 chilometri sotto le catene montuose, la crosta oceanica ha uno spessore compreso tra i cinque e i 10 chilometri e si pensa che sia costituita dallo strato basaltico ricoperto da depositi sedimentari.

Si suppone che sotto la crosta continentale e quella oceanica la discontinuità di Mohorovičić nella velocità delle onde sismiche segni la sommità del mantello. Questa velocità sale bruscamente a 7,8 chilometri al secondo e aumenta in modo irregolare fino a 13,6 chilometri al secondo procedendo verso il basso attraverso il mantello superiore (dai 35 ai 300 chilometri), medio (dai 300 ai 950 chilometri) e inferiore (dai 950 ai 2900 chilometri). Analogamente la densità delle rocce del mantello passa da 3,3 a 5,9 grammi per centimetro cubo. In base a prove dirette, fornite dalle rocce vulcaniche, o indirette, fornite dalle meteoriti e dai dati geofisici e astronomici, si pensa che il mantello sia costituito da silicati di ferro e magnesio fino a 1100 chilometri di profondità. A profondità maggiori prevalgono i solfuri e gli ossidi di ferro, rame, zinco, piombo, mercurio, antimonio e bismuto, come pure selenio, tellurio, oro, argento e altri metalli più pesanti. Nel mantello superiore la temperatura va da 1000 a 1500 gradi centigradi e a quelle profondità la pressione raggiunge le 100 000 atmosfere.

Per verificare e ampliare la raffigurazione della crosta della Terra ottenuta con dati indiretti, i geologi aspirano a eseguire perforazioni che raggiungano la discontinuità di Conrad e, ancora più in basso, quella di Mohorovičić. La Penisola di Kola è stata scelta perché in questa regione lo scudo baltico è rappresentativo delle antiche zolle continentali granitiche di India, America settentrionale, Africa meridionale, Australia occidentale, Antartide e Groenlandia. Il sito si trova inoltre nella regione mineraria di Pechenga dove vi sono giacimenti di rame e di nichel: si sperava che un pozzo in quella zona avrebbe chiarito la genesi di quei giacimenti. La regione, soggetta per centinaia di milioni di anni a glaciazioni e ad alterazione dovuta agli agenti atmosferici, ha perso per erosione 5000-15000 metri della parte superiore, di strato granitico. Perciò la sezione geologica di 12 000 metri attraversata dal pozzo corrisponde a uno strato continentale «medio» che generalmente si trova tra 8000 e 20000 metri al di sotto della superficie.

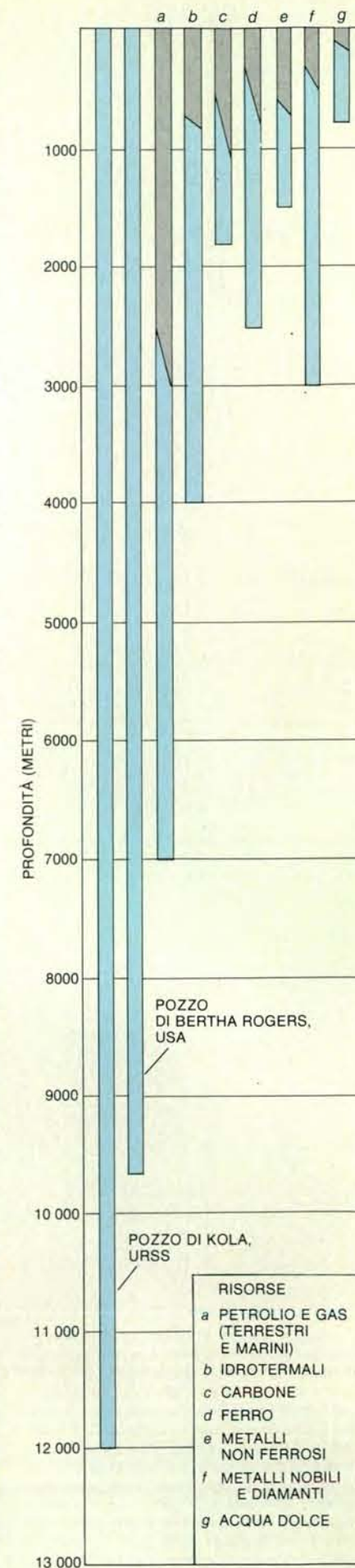
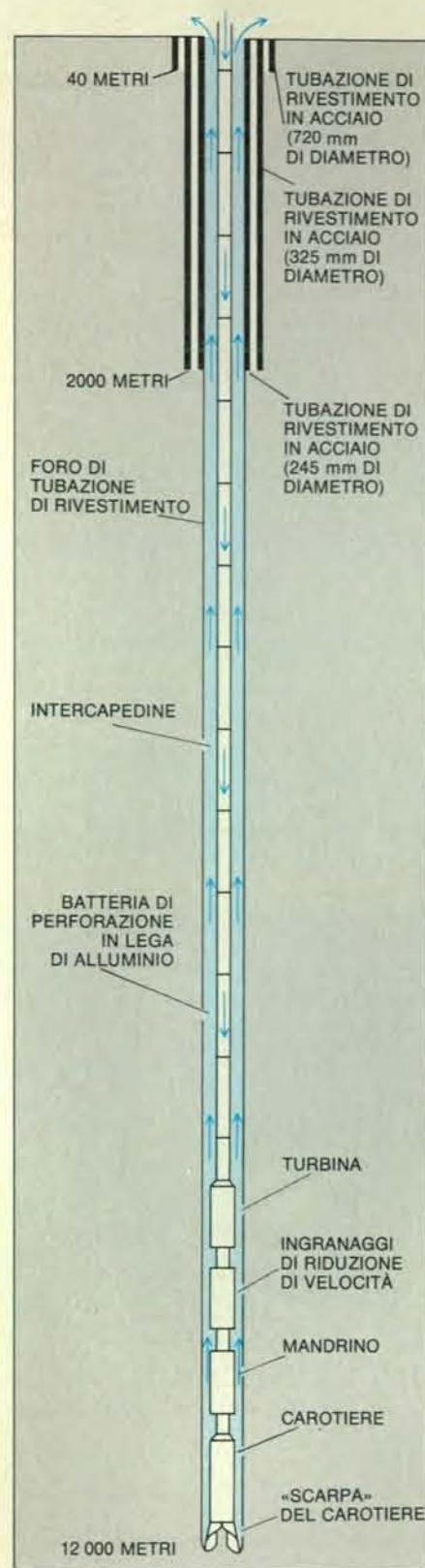
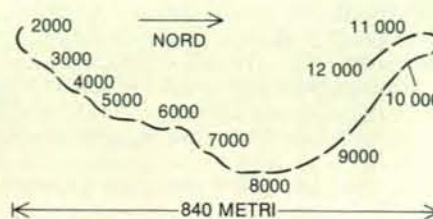
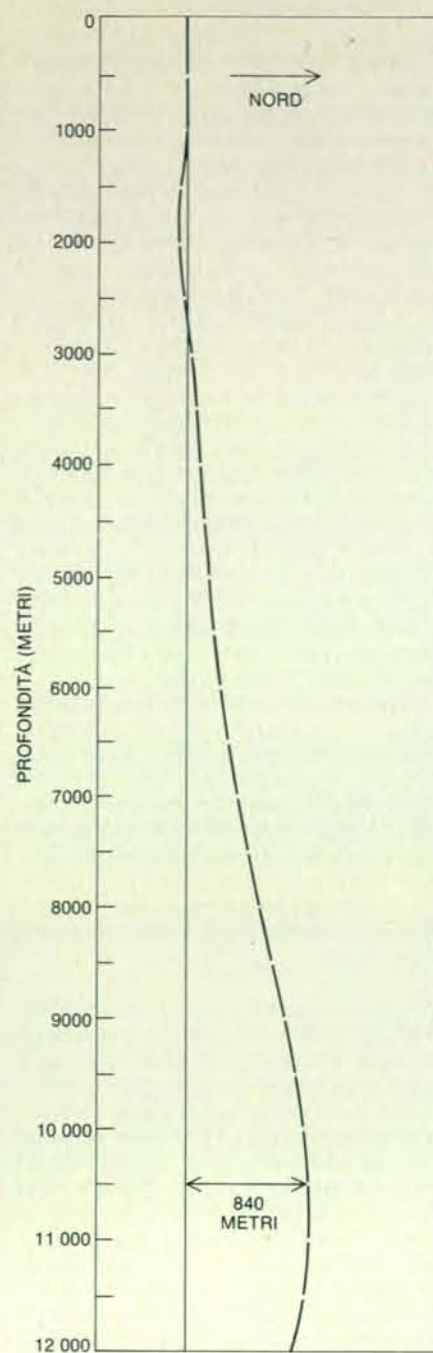
La perforazione del pozzo di Kola ha apportato significative innovazioni alla tecnologia di trivellazione. Sotto i 10000 metri la trivellazione a rotazione convenzionale, che fa girare la punta della trivella ruotando l'intera batteria di perforazione, incontra difficoltà insormontabili. Il peso di 800-900 tonnellate dell'insieme di aste d'acciaio dà origine in superficie a sollecitazioni enormi e amplifica le forze che si oppongono alla rotazione della batteria. Nel pozzo di Kola una turbina nel fondo, azionata dal flusso del fango di perforazione, fa girare lo scalpello. La rotazione della batteria è totalmente eliminata o ridotta a pochi giri al minuto quando occorre potenziare la rotazione dello scalpello ottenuta con la turbina.

In un pozzo comune il fango di perforazione viene pompato dentro la batteria di perforazione per raffreddare lo scalpello e asportare i detriti; risale quindi in superficie attraverso l'intercapedine tra la batteria e la parete del pozzo, che contribuisce a mantenere integra. Nel pozzo di Kola il fango viene pompato verso la turbina con una pressione di 250 atmosfere. Un ingranaggio posto sotto la turbina riduce la velocità di rotazione dello scalpello di tre o quattro volte rispetto a quella delle trivelle dei normali pozzi, fino a ottenere il valore ottimale di 80-150 giri al minuto, e in corrispondenza aumenta il momento di rotazione. Nel sistema di pompaggio una linea idraulica a retroazione, trasmettendo gli sbalzi di pressione del fango di perforazione alla velocità di 1500 metri al secondo, controlla la rotazione della turbina e della batteria e quindi la velocità e il momento della punta della trivella. Filtri antirumore di un modello particolare garantiscono un controllo attendibile a una profondità di 12000 metri.

La batteria di perforazione, in lega di alluminio molto resistente, pesa solo 400-500 tonnellate e ciò è un grande vantaggio perché, per la sostituzione degli scalpelli consumati o per recuperare le carote, è necessario estrarre e far ridiscendere al fondo l'intera batteria centinaia di volte. Per raggiungere gli 11000 metri del pozzo di Kola, la somma delle lunghezze delle aste di trivellazione, moltiplicata per tutte le volte che esse sono passate dalla bocca del pozzo in questi viaggi di andata e ritorno, ha superato i 25 milioni di metri. A prescindere dalla riduzione del carico che deve sostenere la torre di trivellazione, il minor peso della batteria riduce l'usura della batteria stessa, della tubazione del rivestimento e della parete del pozzo, provocata dall'attrito. L'usura può essere notevole perché la perforazione non procede mai perfettamente verticale. Il percorso del pozzo di Kola ha deviato dalla verticale di un valore medio di 5 gradi fino a un massimo di 17 gradi tra 10000 e 10500 metri; a questa profondità lo scostamento laterale dello scalpello rispetto alla testa del pozzo era di 840 metri *(si veda l'illustrazione a destra nella pagina successiva)*.

Il progetto stesso del pozzo di Kola doveva adeguarsi agli imprevedibili cambiamenti di condizioni incontrati a mano a mano che lo scalpello avanzava verso il basso. Mentre i primi due chilometri di pozzo vennero rivestiti *(si veda l'illustrazione a sinistra nella pagina successiva)* fu evidente che a maggiori profondità era necessaria una maggiore flessibilità operativa. Perciò, sotto i 2000 metri, l'intero pozzo fu perforato senza alcun rivestimento.

Ogni campione di roccia estratto dalle profondità della Terra ha il suo valore. A mano a mano che il pozzo di Kola si approfondiva, diventava sempre più difficile portare le carote in superficie. Normalmente una carota di forma cilindrica, con un diametro di 60-80 millimetri, entra nel carotiere a mano a mano che la punta della trivella erode l'anello di roccia del fondo; la carota rimane nel tubo portacarota fino a quando la batteria di perforazione non viene estratta. Tuttavia, a una certa profondità, la roccia «scoppia» per la liberazione di forze interne quando lo scalpello elimina la pressione esercitata dagli strati di roccia sovrastanti, che genera appunto tali forze. Se a questa profondità venisse usata per il carotaggio una normale attrezzatura, i frammenti di carota bloccherebbero l'entrata del tubo portacarota e il 90-95 per cento della carota verrebbe frantumato ed entrerebbe nel fango di perforazione. Un nuovo dispositivo per il prelevamento delle

I pozzi, compresi quelli minerari, variano in profondità a seconda dei giacimenti cercati. Nell'istogramma vengono confrontate le profondità medie *(in grigio)* e massime *(in colore)* raggiunte dai vari tipi di pozzo con quelle del pozzo di Kola e del pozzo di Bertha Rogers nell'Oklahoma. Quest'ultimo, che è il secondo pozzo del mondo in ordine di profondità, serviva a estrarre gas, ma la perforazione venne arrestata a 9674 metri di profondità quando la punta della trivella incontrò zolfo allo stato fuso.

Lo schema mostra il progressivo restringimento del pozzo. Fino a 40 metri fu impiegato uno scalpello con un diametro di quasi un metro e la parete fu rivestita con una tubazione di 720 millimetri di diametro. La perforazione proseguì con uno scalpello di 214 millimetri di diametro. I primi 2000 metri furono poi allargati e rivestiti con una tubazione di 325 millimetri di diametro. A maggior profondità il foro non è rivestito, ma sostenuto dalla pressione del fango di perforazione che scorre verso il basso nelle aste di perforazione e risale in superficie tra aste e parete.

La traiettoria del pozzo, vista di profilo *(in alto)* e come deviazione orizzontale rispetto alla testa del pozzo *(in basso)*, si allontana dalla verticale per un massimo di 840 metri a una profondità di 10 500 metri. Una certa deviazione del foro di perforazione è inevitabile, ma aumenta l'attrito quando la batteria di perforazione viene calata nel foro o viene fatta risalire per recuperare le carote o per sostituire gli scalpelli consumati. Sono stati utilizzati sensori dotati di giroscopi e di fili a piombo per misurare l'inclinazione del foro e per manovrare lo scalpello allo scopo di cercare di ridurre la deviazione.

carote, che devia una parte del flusso di fango nel tubo portacarota, raccoglie i frammenti della carota «esplosa» e li trasporta in una speciale «camera», mantenendo così libera l'entrata per altri campioni.

La tecnologia dimostratasi valida nel pozzo di Kola apre la possibilità di effettuare perforazioni fino a 15 000-17 000 metri. Tuttavia vanno presi in considerazione altri problemi. Le normali leghe di alluminio perdono la loro resistenza a temperature comprese tra i 110 e i 150 gradi centigradi. Noi siamo riusciti a produrre leghe che resistono fino a temperature comprese tra 230 e 250 gradi centigradi. Si prevede che la metallurgia delle polveri consentirà di realizzare leghe di alluminio stabili fino a una temperatura compresa tra 270 e 300 gradi centigradi, le quali devono venire preparate mano a mano che aumenta la profondità dei pozzi. Per resistere a temperature di 400 gradi, riscontrabili a profondità ancora maggiori, si possono costruire aste di perforazione in lega di titanio.

Oltre a risolvere i problemi connessi alle temperature elevate, la ricerca tecnologica per la perforazione profonda deve anche prendere in considerazione pressioni che raggiungono le 3000 atmosfere, gli effetti chimici corrosivi di acque fortemente mineralizzate intrappolate nella roccia, la perdita di stabilità della massa rocciosa intorno al pozzo e la deviazione del pozzo dalla verticale. Il recupero delle carote diventa sempre più problematico. Per risolvere questo problema è necessario progettare un sistema di carotaggio e di trasporto in superficie dei campioni di roccia in contenitori a chiusura ermetica che mantengano le condizioni esistenti al fondo del pozzo, compresa la saturazione di acqua e di gas della roccia. Attualmente si sta preparando in laboratorio una camera pressurizzata per simulare le condizioni esistenti a 15 000-20 000 metri di profondità, cioè temperature comprese tra 300 e 400 gradi centigradi e pressioni comprese tra 2000 e 3000 atmosfere.

La «storia della Terra» rivelata dal pozzo di Kola va letta, naturalmente, dal basso verso l'alto. Nel complesso dell'Archeano, che va dai 12 000 ai 6842 metri, il primo stadio vede l'accumulo di spessi strati sedimentari dovuti all'alterazione da agenti atmosferici dei graniti originari, ripetutamente intercalati da intrusioni di graniti plutonici. Che questi graniti fossero ricchi di ferro e di titanio è provato dalla concentrazione di minerali di magnetite e di ilmenite, che raggiunge il 40-50 per cento della roccia a 8711 metri di profondità. Nel secondo stadio la roccia è stata sottoposta a ripiegamenti, metamorfismo e ultrametamorfismo a temperature comprese tra 750 e 900 gradi centigradi e a pressioni comprese tra 5000 e 11 000 atmosfere.

Poiché le rocce registrano con elevata sensibilità temperatura e pressione, i geologi possono ricostruire attraverso queste la loro storia. Dallo stesso materiale proveniente dal mantello, le rocce metamorfiche sviluppano una grande varietà di caratteristiche distintive, ossia «facies», che possono includere composizioni chimiche elementari diverse a seconda della pressione e della temperatura al momento della loro formazione.

In generale, il metamorfismo dà origine a rocce più dense, con un contenuto minore di acqua chimicamente legata, a partire da rocce più idrate. Gli elementi che non entrano a far parte delle nuove fasi cristalline si disciolgono nell'acqua che viene liberata.

Nella Penisola di Kola la datazione con il carbonio radioattivo fissa il culmine del metamorfismo dell'Archeano tra 2,7 e 2,8 miliardi di anni fa; seguirono una profonda erosione provocata dall'acqua e un accumulo di sedimenti derivati dall'alterazione della crosta, dovuta agli agenti atmosferici, in depressioni isolate. In alcune regioni, soprattutto in Sud Africa, immensi depositi di conglomerati metalliferi sono associati a tali sedimenti.

Il complesso del Proterozoico, che si estende da 6842 metri di profondità alla superficie, ha iniziato a formarsi sul basamento dell'Archeano 1,1 miliardi di anni fa. In questo arco di tempo le rocce registrano quattro fasi principali nella formazione della crosta continentale. Durante la prima fase, sul basamento dell'Archeano si depositò materiale vulcanico sedimentario. Gli strati ghiaiosi mostrano bruschi cambiamenti di spessore, indice di una deposizione fluviale in antiche valli. Seguirono due cicli di plutonismo, il primo dei quali provocò l'intrusione di corpi granitici, privi di elementi metallici, nelle rocce precedentemente formatesi, che subirono di conseguenza un metamorfismo a bassa temperatura. Nel secondo ciclo il mantello fornì rocce ricche di elementi metallici. Queste intrusioni metallifere diedero luogo ai giacimenti di solfuro di rame e di nichel che affiorano nella regione di Pechenga. Nel pozzo di Kola questi depositi sono stati trovati a intervalli, fino a una profondità di 1500-1800 metri. Nella quarta fase ebbe luogo l'anomalo episodio di metamorfismo «chiuso» che portò alla disgregazione idraulica delle rocce metamorfiche osservata per la prima volta nel pozzo di Kola, nella zona spessa 4500 metri che attraversa il basamento dell'Archeano.

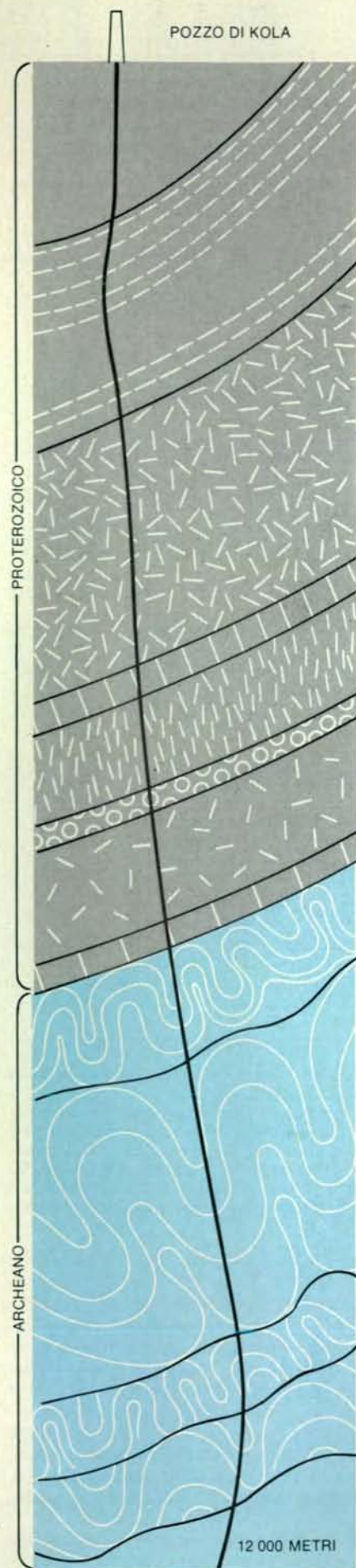
La sezione trasversale del pozzo cambia con la profondità, sebbene il diametro dello scalpello sia rimasto costante. L'istogramma *(a destra)* riporta il calibro del foro a diverse profondità riferito all'asse maggiore; la curva rappresenta l'andamento medio dei valori. La forma generalmente ellittica del foro è dovuta a forze compressive diseguali che agiscono nella roccia orizzontalmente; l'asse minore del foro corrisponde alla direzione in cui la forza è maggiore. La roccia si è frantumata ed è «scoppiata» lungo l'asse di minore forza, ampliando di conseguenza il foro e producendo frammenti che sono stati poi portati in superficie dal fango di perforazione. La dimensione dell'effetto prodotto è stata diversa a seconda della pressione esercitata e del tipo di roccia.

Le carote mostrano che il contenuto di acqua chimicamente legata rimane costante, intorno al 4 per cento, fino a 4500 metri di profondità. Esso diminuisce quindi piuttosto bruscamente fino al 2,1 per cento. A questa profondità inizia la zona di disgregazione idraulica in cui la microfratturazione provoca un aumento della porosità della roccia di tre o quattro volte rispetto alla porosità osservata nelle rocce sovrastanti e, di conseguenza, una diminuzione della densità della massa rocciosa da 3,1 a 2,9 grammi per centimetro cubo. Si calcola che l'acqua liberata, intrappolata negli interstizi della roccia fratturata, abbia fatto aumentare dell'1,7 per cento il volume totale iniziale occupato dalla roccia e dall'acqua. L'enorme pressione idraulica esercitata ha causato così la microfratturazione della roccia che inizialmente deve aver determinato un aumento della porosità di 10 volte rispetto a quella degli strati sovrastanti.

Il limite inferiore di questa zona, a 9000 metri, è segnato da un aumento di velocità delle onde sismiche, che non evidenzia, quindi, la prevista discontinuità di Conrad tra il granito e il basalto, ma segna semplicemente la fine della zona di disgregazione con il ritorno a rocce di densità normale e la cessazione dell'afflusso di acque termali nel pozzo.

Il valore di uno sguardo diretto all'interno della Terra è ampiamente dimostrato da questa scoperta. La disgregazione idraulica delle rocce metamorfiche può spiegare la natura geologica delle discontinuità evidenziate dai cambiamenti di velocità e di riflessione delle onde elastiche osservati in tutto il mondo nella parte superiore della crosta (fino a 15-20 chilometri di profondità). Inoltre questo fenomeno porta un cambiamento sostanziale nelle teorie sulla circolazione dell'acqua nella crosta continentale e sulla natura dell'idrosfera sotterranea.

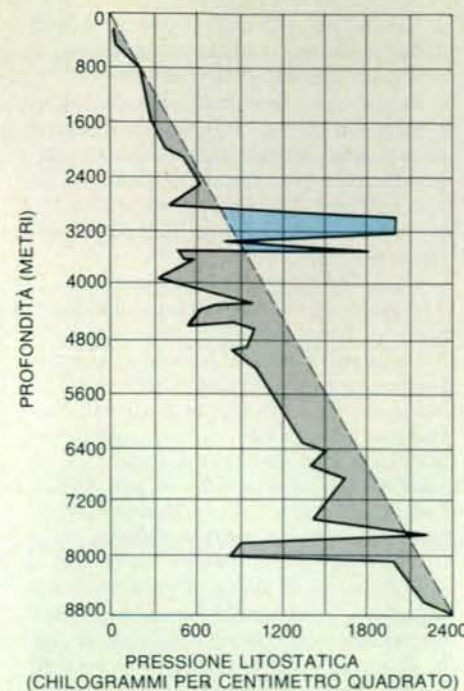
L'elevato contenuto di minerali delle acque interstiziali e la presenza di gas nel pozzo dimostrano che nelle rocce cristal-

line a grandi profondità hanno luogo attivi processi tra gas e acqua. Di conseguenza aumentano le possibilità di scoprire nuovi depositi di minerali a quelle profondità. Nella zona di fratturazione che si incontra nel pozzo di Kola i frammenti di roccia sono cementati fra loro da solfuri di rame, nichel, ferro, zinco e cobalto. La temperatura di formazione relativamente bassa di questi solfuri e la composizione isotopica dello zolfo, che assomiglia a quella delle meteoriti, suggeriscono che questi solfuri abbiano avuto origine nel mantello. Quindi in uno spessore notevole di crosta continentale vi sono condizioni favorevoli alla formazione di giacimenti minerari idrotermali.

La misurazione diretta delle temperature nel pozzo obbliga a rivedere le teorie sulla distribuzione e sul flusso del calore all'interno della Terra. Si suppone che in una regione tettonicamente stabile, come lo scudo baltico, la temperatura aumenti con la profondità solo lentamente, raggiungendo i 50 gradi centigradi a 7000 metri e forse i 100 gradi a 10 000 metri. In realtà il gradiente termico misurato coincide con l'aumento previsto di un grado ogni 100 metri solo fino a una profondità di 3000 metri; al di sotto, la temperatura inizia ad aumentare di 2,5 gradi ogni 100 metri e a 10 000 metri raggiunge i 180 gradi. Centinaia di metri cubi di fango di perforazione freddo, pompati nel pozzo, sono tornati in superficie riscaldati a 45 gradi. Poiché la radioattività delle rocce attraversate dal pozzo contribuisce solo in modo insignificante a questo flusso termico, esso deve chiaramente provenire dal mantello sottostante.

Il successo ottenuto con il pozzo di Kola ha infuso nuovo ottimismo per quanto riguarda la perforazione sistematica profonda e semiprofonda della crosta terrestre entro i confini dell'Unione Sovietica *(si veda l'illustrazione in basso a pagina 76)*. I siti in cui perforare i pozzi sono stati scelti nei punti di intersezione di una rete di profili sismici, utilizzando i risultati di un vasto programma di prospezione sismica profonda realizzato nel decennio scorso.

La perforazione dei pozzi servirà, a sua volta, a migliorare l'interpretazione dei dati sismici. L'osservazione diretta delle condizioni e lo studio delle carote prelevate dal pozzo serviranno a stabilire correlazioni esatte tra le discontinuità geodinamiche della crosta e del mantello superiore e le discontinuità e i contatti di struttura e di composizione osservati. Questo lavoro servirà anche a migliorare la risoluzione e la precisione delle registrazioni sismiche e quindi ad aumentare la quantità di informazioni che si possono ottenere da osservazioni geofisiche indirette. Le anomalie nella propagazione e nella velocità delle onde sismiche non correlate con dati strutturali raccolti nei pozzi possono rappresentare nuovi e importanti problemi da risolvere.

La perforazione continuerà nei pozzi di Kola e di Saatly, che saranno utilizzati come laboratori per studi intensivi sia sulla dinamica sia sulla struttura della crosta in quei siti. I prossimi tre pozzi superprofondi (più di 7000 metri) verranno perforati a Muruntan, Anastasievsko-Troitskaya e vicino al mar Caspio. Contemporaneamente si procederà alla perforazione di sei pozzi profondi (più di 4000 metri), tre dei quali in zone ricche di idrocarburi e gli altri tre in zone ricche di minerali. In tal modo, oltre a rispondere a fondamentali quesiti sulla struttura della crosta continentale, i profili sismici profondi, combinati con l'osservazione diretta in pozzi profondi e superprofondi, permetteranno di valutare i principali elementi geotettonici dell'Unione Sovietica, ritenuti sedi di importanti risorse.

Gli strati geologici attraversati dal pozzo abbracciano 1,4 miliardi di anni di storia della Terra. Le rocce sedimentarie e vulcaniche trovate fino a 6800 metri di profondità *(in grigio)* risalgono al Proterozoico, che iniziò 2,4 miliardi di anni fa. Gli strati granitici più profondi *(in colore)* si formarono 2,7 miliardi di anni fa nell'Archeano, il periodo più antico del tempo geologico. Si pensava che intorno ai 9000 metri di profondità, dove si registra una brusca variazione di velocità delle onde sismiche, il pozzo passasse da rocce granitiche a un basamento basaltico. Invece si è scoperto che il cambiamento di velocità segna il limite inferiore di una zona, spessa 4500 metri, di rocce disgregate dalla pressione dell'acqua liberata da minerali cristallini in processi metamorfici.

La pressione litostatica, come risulta dalle misurazioni della velocità delle onde acustiche che attraversano la roccia in vicinanza del foro, spesso non presenta un incremento lineare con la profondità *(linea tratteggiata)* come è prevedibile per un materiale omogeneo. La zona di pressione anormalmente alta a 3200 metri di profondità riflette l'elevata densità di strati rocciosi impermeabili. Le pressioni sproporzionalmente basse tra 4000 e 9000 metri corrispondono a una zona di roccia fratturata.

L'immaginazione spaziale

Da una serie di esperimenti condotti per studiare la forma di pensiero coinvolta nelle operazioni spaziali è emerso che la rotazione immaginata di un oggetto rispecchia fedelmente quella che avviene nel mondo fisico

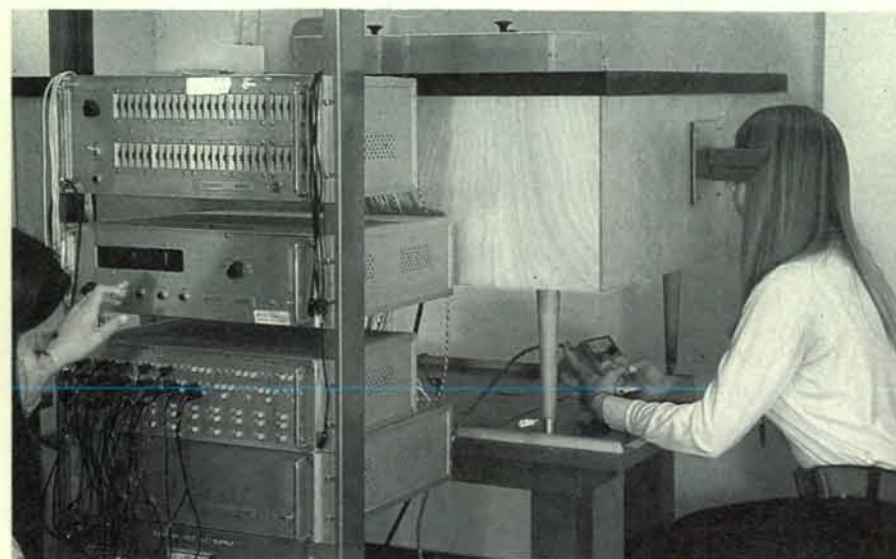
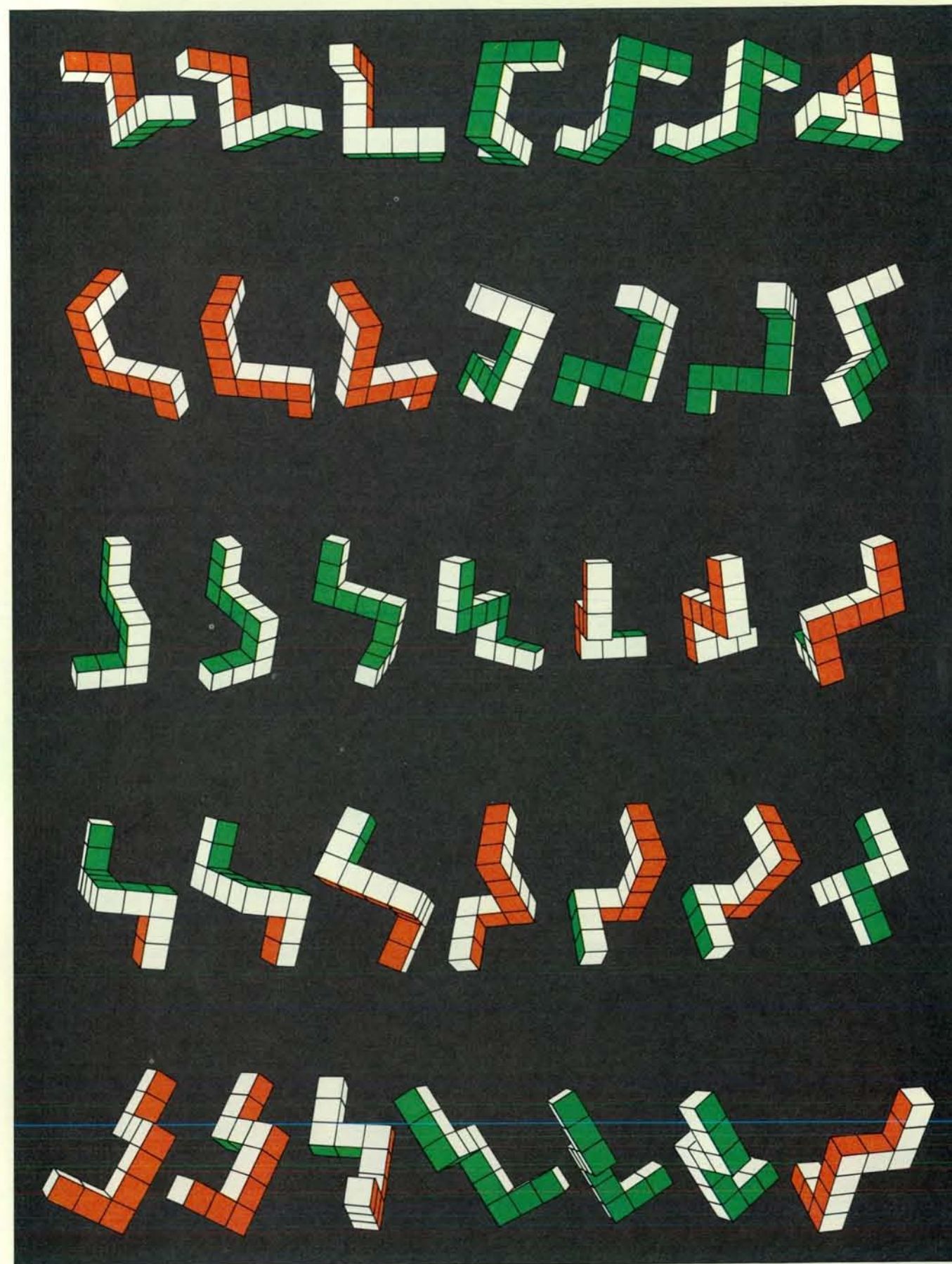
di Lynn A. Cooper e Roger N. Shepard

Che cos'è il pensiero? Una prima risposta ci può venire dall'introspezione. Alcuni pensieri sono verbali: una sorta di colloquio silenzioso con se stessi. Altri processi mentali sembrano di natura visiva: vengono richiamate alla mente immagini, elaborate poi senza l'intervento delle parole. I meccanismi del pensiero, benché siano evidenti, sono sfuggiti a lungo all'analisi sperimentale e alla quantificazione. Come è possibile misurare e studiare scientificamente questi processi soggettivi, che sono apparentemente inaccessibili?

Abbiamo cominciato a dare una risposta a questa domanda, ideando esperimenti per sondare la natura di un tipo particolare di processo di pensiero: l'immaginare operazioni che si svolgono nello spazio. I nostri risultati danno una conferma empirica di quel che è soggettivamente evidente: la mente può rappresentare processi fisici e sottoporli ai vincoli geometrici che vigono nel mondo esterno. Nella vita quotidiana gli esempi di queste operazioni mentali abbondano. Come si fa a far passare un tavolino da gioco attraverso il vano di una porta senza piegarne le gambe? Molti sostengono di dover prima immaginare il processo di girare il tavolo su un fianco, introdurre due gambe del tavolo nel vano della porta e poi girare e spostare il tavolo in modo che le prime due gambe, il piano e le altre due gambe passino attraverso la porta.

Questo tipo di immaginazione spaziale può non essere una caratteristica esclusiva degli esseri umani. Uno di noi ha osservato un cane pastore tedesco recuperare un lungo bastone che era stato gettato oltre uno steccato al quale mancava uno dei paletti verticali. Il cane si tuffò attraverso il varco, prese in bocca il bastone e si diresse di nuovo di corsa verso l'apertura, con il bastone di traverso. Proprio quando la catastrofe sembrava imminente, il cane si arrestò, fece una pausa e ruotò la testa di 90 gradi. Tenendo il bastone verticalmente, passò senza danno attraverso lo steccato. L'operazione che ebbe luogo nella mente del cane prima di girare la testa probabilmente era non verbale. Non potrebbe essersi trattato di una rotazione mentale, preparatoria, del bastone? (E non è stato forse attraverso un processo di visualizzazione spaziale, anziché di deduzione verbale, che voi leggendo vi siete resi conto di quale fosse la catastrofe e di come sia stata evitata?)

La capacità di rappresentare oggetti o gruppi di oggetti disposti in un certo ordine e le loro trasformazioni nello spazio è molto preziosa per fronteggiare la realtà concreta della vita quotidiana, poiché permette di pianificare le proprie azioni e di prefigurarne i risultati. Molti scienziati hanno sostenuto di aver raggiunto i risultati migliori sulla base di relazioni e trasformazioni spaziali che si erano immaginati mentalmente. Due casi famosi sono quello di Friedrich Kekulé con la sua immagine della struttura della molecola del benzene e quello di James Watt con la sua visualizzazione del meccanismo fondamentale della macchina a vapore. Albert Einstein ha perfino asserito di essere arrivato alla teoria della relatività «visualizzando...effetti, conseguenze e possibilità» attraverso «immagini più o meno chiare che si possono "spontaneamente" riprodurre e combinare».

Affermazioni soggettive e qualitative, anche se fatte da scienziati, non possono tuttavia sostituirsi al ragionamento oggettivo e quantitativo. Come disse nel 1923 lo psicofisiologo K. S. Lashley, «l'introspezione può operare l'indagine preliminare, ma poi dev'essere seguita da una serie di misurazioni oggettive». Solo lentamente, però, si è giunti a quest'indagine

Il tachistoscopio è un'apparecchiatura che permette al ricercatore di cronometrare le risposte di un soggetto a stimoli visivi. Quando il ricercatore preme un interruttore, si illumina uno schermo, compaiono le immagini e parte un cronometro. La reazione del soggetto che, nel caso illustrato, viene registrata attraverso comandi manuali, fa fermare il cronometro. Si ottiene in questo modo una registrazione precisa dell'intervallo di tempo intercorso fra lo stimolo e la risposta.

Le forme tridimensionali ruotate nello spazio sono disegni prospettici generati dal calcolatore. Quando ai soggetti venivano mostrate coppie di disegni al tratto raffiguranti la stessa forma diversamente orientata, per riconoscerne l'identità essi impiegavano un tempo proporzionale alla differenza angolare tra le orientazioni. Il fatto che il tempo occorrente per il confronto aumentasse linearmente con la differenza di orientazione ha suggerito che i soggetti dovessero immaginare di ruotare una forma allo stesso modo dell'altra per vedere se coincidevano.

sistematica dell'immaginazione spaziale.

Per tutta la prima metà del nostro secolo, gli ostacoli teorici sono stati fortissimi. In questo periodo gli esponenti del comportamentismo americano, da J. B. Watson a B. F. Skinner, tentarono di separare la psicologia dalle sue origini filosofiche legate all'introspezione e di basarla sui suoi fondamenti empirici. I comportamentisti sostenevano che tutti i termini teorici debbono corrispondere a stimoli e relative risposte agli stimoli, precisabili in modo oggettivo, e che vanno banditi i riferimenti a fenomeni soggettivi insiti in termini come coscienza, mente, pensiero e immaginazione. Le ricerche di laboratorio si concentrarono su eventi fisici registrabili come la pressione esercitata da un ratto su una sbarra o il becchettio di una leva da parte di un piccione, invece che sul lavorio nascosto della mente.

Dopo la metà del secolo, nuove scoperte cominciarono a smantellare le barriere che i comportamentisti avevano eretto contro lo studio delle strutture e dei processi del pensiero. Il linguista Noam Chomsky ha portato solide argomentazioni a favore dell'idea che il comportamento linguistico sia guidato da schemi innati che erano stati completamente trascurati dai comportamentisti. Gli psicologi sperimentali, poi, hanno fornito dimostrazioni convincenti della possibilità di dedurre, e perfino quantificare, i processi mentali a partire dalle configurazioni dei dati registrati in modo oggettivo.

Le viste prospettiche mostrate a coppie ai soggetti del primo esperimento condotto dagli autori presentavano tre tipi di differenze. Nel primo caso (*in alto*) i disegni raffiguravano oggetti identici in posizioni che differivano di una rotazione nel piano dell'immagine. Nel secondo (*al centro*) le posizioni differivano per una rotazione nella terza dimensione. I soggetti individuavano l'identità di entrambi i tipi di oggetti presentati a coppie altrettanto rapidamente, il che fa pensare che in ambedue i casi, per poterli confrontare, immaginassero gli oggetti come solidi tridimensionali rotanti nello spazio. Nel terzo caso si trattava di forme enantiomorfe, ossia speculari (*in basso*).

Gli esperimenti che abbiamo progettato per sondare quel tipo di processi mentali che i comportamentisti avevano ignorato sono stati costruiti in modo da soddisfare anche la loro esigenza di lavorare su dati oggettivi e quantitativi. La prima delle nostre ricerche sul processo di immaginazione spaziale, intrapresa nel 1971 da uno di noi (Shepard) insieme con Jacqueline Metzler, soddisfaceva ambedue i criteri: ogni prova sperimentale era oggettiva nel senso che la risposta del soggetto a uno stimolo era oggettivamente giusta o sbagliata, ed era quantitativa nel senso che la variabile presa in considerazione era il tempo impiegato dal soggetto per dare la risposta corretta.

I soggetti dell'esperimento dovevano confrontare una coppia di disegni prospettici generati dal calcolatore. Ciascun disegno raffigurava un oggetto tridimensionale composto da 10 cubetti accostati l'uno all'altro in modo da formare una struttura a braccio con tre piegature ad angolo retto (*si veda l'illustrazione nella pagina precedente*). Alcune coppie mostravano strutture identiche, ma di solito con orientazioni spaziali diverse; altre, frammiste in modo casuale, nella serie di prove, a quelle del primo tipo, mostravano strutture differenti sia per forma sia per orientazione: si trattava in questo caso di strutture enantiomorfe, generate l'una dall'altra per riflessione nello spazio, e diverse dunque fra loro più o meno come la mano sinistra dalla destra.

A ogni prova il soggetto guardava in un tachistoscopio, una scatola buia nella quale si possono mostrare, per tempi esattamente controllabili, stimoli visivi sistemati dietro la scatola. Lo sperimentatore inseriva una coppia di disegni e girava un interruttore per illuminarli, facendo partire contemporaneamente un cronometro. Il soggetto confrontava i disegni il più rapidamente possibile e rispondeva abbassando una delle due leve a sua disposizione: la destra per le coppie della stessa forma e la sinistra per quelle con forme diverse. Quando il soggetto rispondeva, il cronometro veniva fermato e si registrava il tempo impiegato per il confronto.

Poiché le coppie di disegni rappresentavano oggetti che erano identici o differivano per una riflessione nello spazio, i soggetti non potevano basarsi, nel confronto, sull'aspetto superficiale dello stimolo. Il numero di cubetti tra due piegature successive, per esempio, era identico nei due disegni, indipendentemente dal fatto che le strutture fossero identiche o enantiomorfe. Era quindi da escludere una ricerca veloce (una sorta di scorciatoia) di differenze ovvie. I soggetti riferirono di poter confrontare le forme solo immaginando di ruotare uno dei due oggetti fino a fargli raggiungere l'orientazione dell'altro, per poi vedere se coincidevano. In genere - dissero - immaginavano di ruotare l'oggetto di sinistra in modo da collocarne il braccio superiore parallelo a quello corrispondente dell'oggetto di destra; poi cercavano mentalmente di vedere se la parte all'altra estremità dell'oggetto si proiettava nella stessa direzione della parte analoga della seconda struttura.

I tempi di reazione, calcolati a partire dal momento in cui venivano mostrati i disegni fino a quello in cui ciascun soggetto dava un risposta abbassando la leva, forniscono una prova oggettiva a sostegno dei resoconti soggettivi. I tempi aumentavano secondo una funzione lineare della differenza angolare fra le orientazioni dei disegni. Quando venivano mostrati oggetti uguali orientati nello stesso modo, i soggetti impiegavano circa un secondo per riconoscere che erano uguali; al crescere della differenza angolare, i tempi di risposta si innalzavano regolarmente fino a una media di 4,4 secondi per la massima differenza angolare possibile di 180 gradi. Tutti gli otto giovani che presero parte all'esperimento fecero registrare un incremento lineare nei tempi di reazione, ma la pendenza della funzione variava da un individuo all'altro.

L'incremento lineare fa pensare che i soggetti confrontassero gli oggetti immaginando di ruotarne uno fino a orientarlo come l'altro, con una velocità regolare che raggiungeva i 180 gradi in 3,4 secondi in media (4,4 meno il secondo necessario a confrontare due oggetti aventi la stessa orientazione), vale a dire con una velocità media di 53 gradi circa al secondo. È possibile escogitare anche altri metodi per discriminare oggetti identici ed enantiomorfi, ma nessuno richiederebbe un tempo proporzionale alla differenza angolare. Per esempio, è da escludere che i soggetti analizzassero separatamente ciascun disegno per cercare di ricondurne la struttura a un codice di un certo tipo e poi confrontare le descrizioni in codice. Un codice potrebbe consistere nella descrizione del numero di cubetti e della direzione delle piegature successive a partire da un'estremità dell'oggetto: 2D2A2S1 starebbe per «due cubetti, piegatura a destra, due cubetti, piegatura verso l'alto, due cubetti, piegatura a sinistra, un cubetto». Il tempo necessario per generare questo codice dipenderebbe da come ciascun oggetto è orientato. Dato però che i codici vengono calcolati indipendentemente l'uno dall'altro, la somma dei tempi occorrenti per generare i codici per i due oggetti non varia necessariamente al variare della loro differenza angolare.

I risultati non solo indicano nella rotazione mentale la base di questo tipo di confronto, ma chiariscono anche come le immagini mentali dei soggetti riflettessero la struttura tridimensionale degli oggetti raffigurati e non semplicemente le caratteristiche bidimensionali del disegno. Metà delle coppie con la stessa forma differivano per una rotazione nel piano bidimensionale del disegno: i due disegni allora erano identici, tranne che per la posizione nel piano. L'altra metà delle coppie raffigurava oggetti che differivano per una rotazione nella terza dimensione, ossia in profondità (*si veda l'illustrazione nella pagina a fronte*). Benché gli oggetti tridimensionali raffigurati da ogni coppia del secondo insieme di disegni fossero identici, i disegni stessi, in quanto configurazioni bidimensionali di linee e angoli, differivano spesso notevolmente fra loro: una rotazione in profondità trasferisce alcune caratteristiche di un oggetto tridimensionale nel piano del disegno e contemporaneamente ne annulla altre.

Tuttavia la pendenza della funzione tempo di reazione per le coppie che corrispondevano a una rotazione nella terza dimensione non era maggiore che per le coppie in cui la differenza fra i disegni derivava da una rotazione nel piano del disegno. La velocità con cui veniva compiuta mentalmente la rotazione era sempre la medesima, sia quando si trattava di una trasformazione nelle tre dimensioni sia quando la rotazione sembrava aver luogo nelle due dimensioni. I risultati concordavano con quanto riferivano i soggetti, che sostenevano di interpretare i disegni come solidi nello spazio tridimensionale, qualunque fosse l'orientazione dell'uno rispetto all'altro. I soggetti, pertanto, trovavano tutte le rotazioni egualmente facili da immaginare.

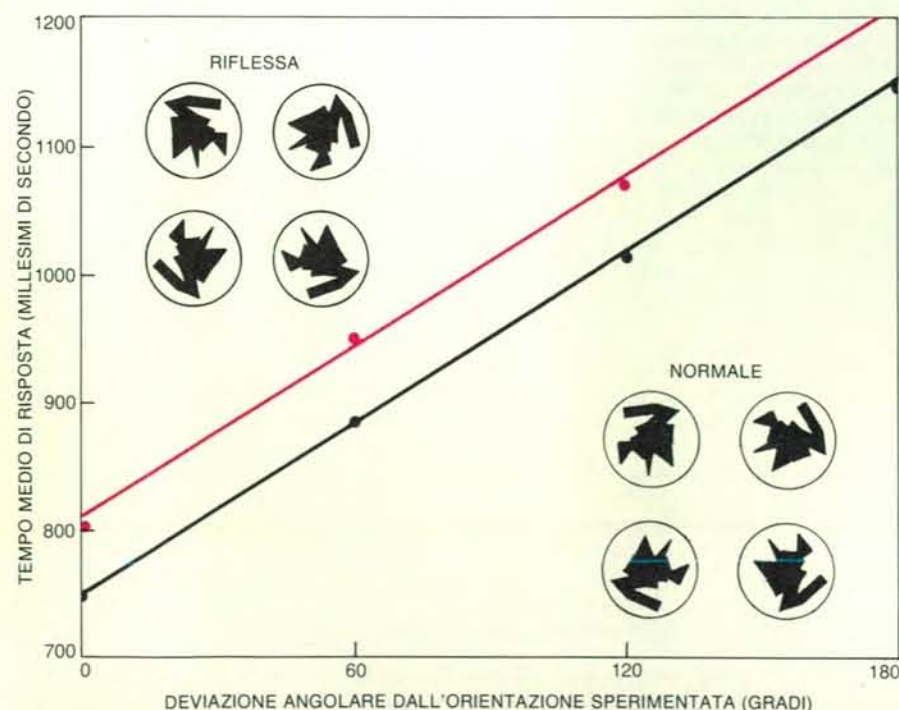
Il carattere progressivo e spaziale del processo mentale di rotazione, stabilito dal primo esperimento, fa pensare a un'analogia con le trasformazioni del mondo naturale. Si è tentati di considerare la rotazione mentale come la simulazione interiore di una rotazione esterna. Una descrizione in questi termini, tuttavia, sarebbe giustificata solo se potessimo dimostrare che i processi interiori passano attraverso stadi intermedi che corrispondono alle posizioni intermedie di un oggetto fisico che ruoti nel mondo esterno.

Per ottenere questa prova ulteriore in merito alla natura analogica della rotazione mentale, uno di noi (Cooper) compì una serie di esperimenti in cui i soggetti dovevano reagire non a una coppia di oggetti, bensì a una singola figura presentata a intervalli e con orientazioni diverse. Facendo scattare il processo di immaginazione di una rotazione con un singolo stimolo e poi presentando lo stesso oggetto, o uno diverso, orientato in un modo qualsiasi e dopo un intervallo qualunque, potevamo sondare la trasformazione mentale mentre aveva luogo.

Tuttavia, prima di studiare il processo di rotazione nel suo svolgimento, era necessario risolvere un problema iniziale, che richiese una breve digressione sperimentale. Dovevamo accertare se la tecnica dello stimolo singolo portasse i soggetti a immaginare le rotazioni mentali progressive proprio come avveniva nel primo esperimento con le coppie di disegni. Per valutare il procedimento con un singolo stimolo chiedemmo ancora una volta ai soggetti di distinguere tra un oggetto e la sua immagine speculare, ma in questo caso usammo figure piane bidimensionali anziché solidi tridimensionali come nel primo esperimento: di conseguenza tutte le posizioni mostrate differivano per rotazioni in un piano.

I soggetti appresero dapprima a distinguere la versione «normale» da quella speculare («riflessa») di otto poligoni, la cui orientazione venne tenuta costante nel corso dell'addestramento (*si veda l'illustrazione in alto nella pagina successiva*). Una volta che essi ebbero appreso queste otto distinzioni, cominciò l'esperimento vero e proprio.

Quando i soggetti dovevano confrontare coppie di disegni raffiguranti oggetti identici, ma orientati in modo diverso, i tempi medi di risposta erano direttamente proporzionali alla differenza angolare fra le orientazioni degli oggetti. La relazione lineare ha fatto pensare che il confronto presupponesse un processo di rotazione mentale. La pendenza della funzione, da cui si ricava la velocità della rotazione immaginata, non era maggiore per le orientazioni che differivano per una rotazione in profondità rispetto a quelle che differivano per una rotazione nel piano dell'immagine.

Il gruppo di forme bidimensionali comprende otto poligoni diversi e le loro immagini speculari. Prima di un insieme di esperimenti in cui ai soggetti veniva mostrata una delle forme in una orientazione insolita, ogni partecipante imparava a distinguere la versione «normale» (*a sinistra*) di ciascuna forma orientata in un dato modo da quella «riflessa» (*a destra*). Le otto forme differiscono per il numero di punti che ne definiscono il perimetro, una caratteristica che forse corrisponde alla complessità psicologica. La complessità di una forma, tuttavia, non ha avuto influenza sulla velocità con cui i soggetti hanno distinto la sua versione normale da quella riflessa.

Quando ai soggetti veniva mostrato un poligono noto in una nuova orientazione ed essi indicavano se era la versione normale (*parte inferiore del grafico*) o quella riflessa (*parte superiore del grafico*) emergeva una relazione lineare tra il tempo della loro risposta e l'orientazione del poligono. Per valutare la forma i soggetti evidentemente dovevano immaginare una rotazione della forma normale finché corrispondeva o no a quella mostrata. Se non vi era coincidenza, a essi occorreva ulteriore tempo per indicare che quella che avevano di fronte era la versione riflessa.

In ciascuna prova veniva mostrato a un soggetto uno dei poligoni, orientato come nell'esperimento preliminare o con un'orientazione differente per un multiplo di 60 gradi. Al soggetto veniva richiesto di determinare se si trattasse della versione normale oppure di quella riflessa; per operare la distinzione il soggetto, presumibilmente, doveva immaginare di ruotare il poligono fino a che corrispondesse o no alla rappresentazione mentale della forma normale che aveva appreso nella fase di addestramento. Se corrispondeva, il soggetto premeva un pulsante alla sua destra, per indicare «normale»; in caso contrario, premeva un pulsante alla sua sinistra per indicare «riflessa».

Come già nel caso del procedimento con coppie di stimoli, anche con questo esperimento abbiamo potuto registrare un incremento lineare del tempo di reazione al crescere della deviazione angolare - deviazione, in questo caso, da una posizione nota anziché dalla posizione di un oggetto mostrato contemporaneamente. I soggetti reagirono regolarmente alle versioni normali dei poligoni 60 millisecondi prima che nel caso delle versioni riflesse. Risultò che i soggetti dapprima confrontavano la forma trasformata con quella normale memorizzata ed erano in grado di reagire immediatamente in caso di coincidenza, mentre avevano bisogno di un intervallo di tempo in più per innescare la risposta «riflessa» nel caso in cui le forme non coincidessero.

Le pendenze delle funzioni tempo di risposta erano identiche sia per le forme normali sia per quelle riflesse. La velocità media che si riusciva a dedurre per la rotazione mentale era di 450 gradi al secondo, notevolmente superiore ai 53 gradi al secondo calcolati per l'esperimento precedente con i due stimoli. Evidentemente l'uso di figure piane presentate una alla volta permetteva risposte più rapide, dando ai soggetti la possibilità di concentrare la loro attenzione su uno stimolo singolo. Nonostante le differenze di tipo quantitativo, i risultati di questo procedimento vanno di pari passo con quelli del procedimento precedente.

Dopo aver stabilito che l'esperimento bidimensionale con uno stimolo singolo richiede una rotazione mentale come quella occorrente per il confronto di due stimoli, abbiamo potuto modificare il procedimento allo scopo di studiare più da vicino l'ipotesi che la rotazione mentale sia analoga alle rotazioni del mondo fisico. Il secondo degli esperimenti con uno stimolo singolo verificava direttamente le affermazioni dei soggetti, secondo le quali essi valutavano uno stimolo immaginando di ruotarlo finché non coincideva con la forma appresa. In effetti abbiamo rovesciato la precedente sequenza di eventi: invece di presentare ai soggetti uno stimolo da confrontare con una forma appresa, richiedendo loro così di immaginare una rotazione correttiva, abbiamo chiesto loro per prima cosa di immaginare la rotazione di una forma appresa e solo successivamente abbiamo presentato loro lo stimolo da confrontare.

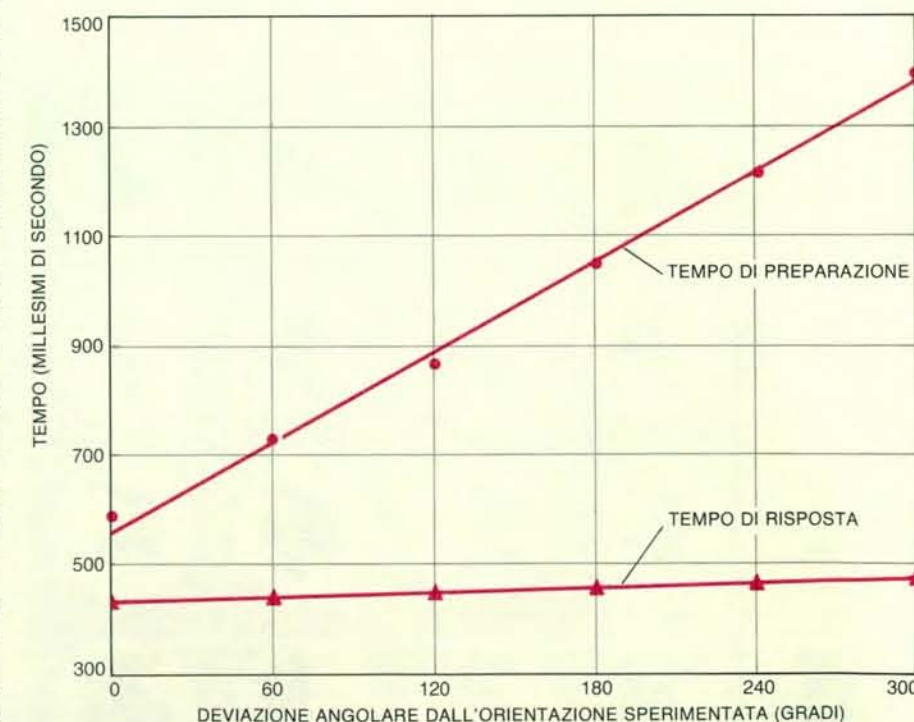
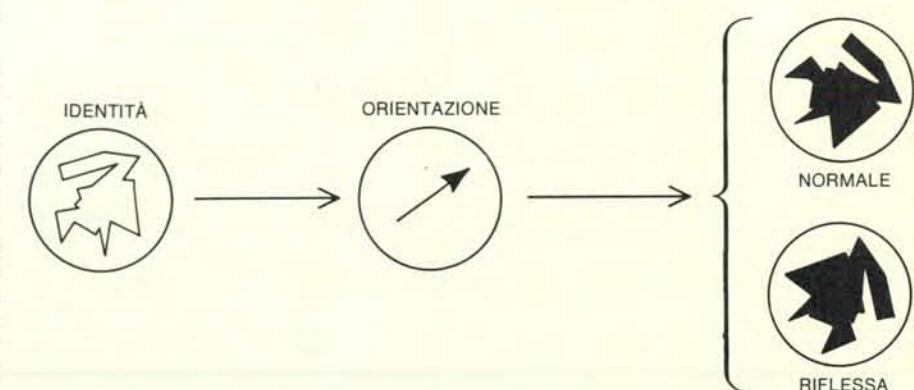
Tutti i soggetti di questo esperimento avevano preso parte al precedente esperimento con uno stimolo singolo e quindi le forme erano loro familiari. In ogni prova il soggetto prendeva visione della sagoma di uno degli otto poligoni normali, che gli veniva mostrata nella stessa orientazione della fase di addestramento. Alla sagoma seguiva un cerchio contenente un indicatore posizionato secondo uno dei sei incrementi angolari posti a intervalli regolari a partire da zero fino a 300 gradi di rotazione in senso orario. Il soggetto allora immaginava di ruotare la sagoma nella direzione che veniva segnalata dall'indicatore. In metà delle prove ai soggetti era stato richiesto di compiere una rotazione mentale in senso orario; nell'altra metà di immaginare di ruotare la forma che veniva loro mostrata in senso antiorario.

Per indicare che l'operazione mentale era compiuta, il soggetto premeva un pulsante, facendo così arrestare il cronometro che registrava il tempo di preparazione. Contemporaneamente compariva una versione del poligono orientata come indicato dalla freccia e partiva un secondo cronometro. Il più rapidamente possibile il soggetto indicava se il disegno rappresentava la versione normale oppure la versione riflessa della forma, pronunciando «N» (per «normale») o «R» (per «riflessa») in un microfono. Un relé attivato dalla voce faceva allora fermare il secondo cronometro.

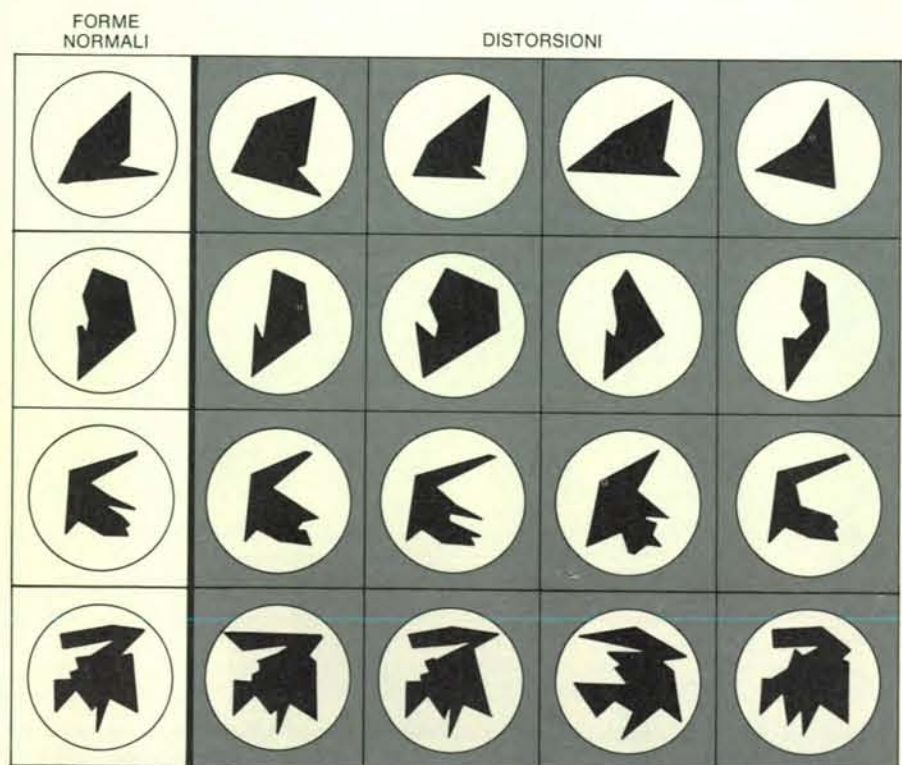
Ciascuna prova dava, quindi, due tempi: quello necessario per effettuare la rotazione mentale e l'intervallo occorrente al soggetto per classificare lo stimolo. Tutti e due gli insiemi di risultati confermarono quanto si era precedentemente scoperto. Il tempo necessario per la rotazione mentale di preparazione cresceva linearmente con la deviazione angolare dalla orientazione di addestramento, come i risultati ottenuti in precedenza facevano prevedere. Inoltre negli esperimenti precedenti non era specificata la direzione della rotazione mentale e la rotazione massima verificabile era di 180 gradi. In questo caso l'incremento lineare nel tempo di preparazione arrivava fino alla rotazione oraria o antioraria massima di 300 gradi, fornendo una prova ulteriore dell'analogia tra operazione mentale e rotazione fisica. La velocità desunta della rotazione preparatoria, in media 370 gradi al secondo, era paragonabile ai 450 gradi al secondo calcolati dal precedente esperimento con un singolo stimolo.

Il secondo gruppo di tempi, cioè quelli necessari ai soggetti per reagire allo stimolo che appariva quando avevano segnalato di essere pronti, confermò che l'atto della rotazione mentale li preparava effettivamente a compiere la distinzione. In media classificavano ciascuna forma come normale o come riflessa in meno di mezzo secondo, indipendentemente dalla sua deviazione angolare dalla posizione appresa. Se i soggetti, dopo aver confrontato lo stimolo, avessero avuto bisogno di compiere ulteriori operazioni mentali, probabilmente i tempi di risposta sarebbero aumentati al crescere della deviazione angolare, così come era successo negli esperimenti precedenti in cui non veniva data la possibilità di compiere alcuna rotazione mentale di preparazione.

Gli esperimenti fin qui descritti documentano le caratteristiche di rotazioni mentali complete. Abbiamo scoperto che il tempo necessario cresce in rapporto diretto con l'angolo di rotazione e abbiamo avuto la conferma del fatto che, dopo aver immaginato una forma ruotata nella direzione indicata da uno stimolo fisico, un soggetto può stabilire identità o differenza con velocità uniforme, qualunque sia il grado di rotazione. Per poter dire, però, che il processo mentale è analogo alle rotazioni fisiche dovevamo ancora dimostrare che esso passa attraverso stadi che corrispondono agli angoli intermedi di una rotazione fisica. Se una tale corrispondenza effettivamente esiste, l'angolo, secondo il quale una forma mostrata solleciterà la risposta più rapida da parte di un soggetto che sta immaginando la sua rotazione, dovrebbe aumentare regolarmente e progressivamente con il tempo, di pari passo con la rotazione mentale.

Per verificare questa ipotesi fu ideato un ulteriore esperimento i cui soggetti erano tutti veterani dei due precedenti e

Nell'esperimento rappresentato in alto sono stati misurati i tempi della rotazione preparatoria. Ai soggetti venivano presentate prima la sagoma di uno degli otto poligoni normali e poi una freccia. Essi immaginavano di ruotare la forma, che era presente soltanto nella loro memoria, nella direzione indicata da quella freccia. Quando avevano completato l'operazione mentale, lo segnalavano premendo un pulsante che faceva comparire una forma orientata nel modo indicato. A questo punto dovevano determinare se la forma era in versione normale o riflessa. Il tempo occorrente per la rotazione mentale preparatoria (*curva superiore*) cresceva linearmente con l'angolo; il fatto che la risposta fosse costantemente rapida (*curva inferiore*) ha confermato che il processo di rotazione mentale preparava realmente i soggetti a valutare la forma riorientata.

di conseguenza conoscevano bene gli otto poligoni. A ogni prova appariva nel tachistoscopio un campo circolare vuoto e al soggetto veniva richiesto di immaginare, alla velocità che gli riusciva naturale, un poligono dato che ruotava all'interno del campo. Dopo un intervallo di tempo non precisato, compariva quella forma o la sua immagine speculare; il soggetto, allora, il più rapidamente possibile, identificava di

L'esperimento di verifica delle rotazioni mentali in atto ha portato ai risultati visualizzati nel grafico. Quando un soggetto immaginava di far ruotare una forma e la presentazione di una forma di prova veniva sincronizzata in modo che la sua orientazione corrispondesse a quella dell'immagine mentale in quel momento, i tempi di risposta erano costantemente rapidi per qualsiasi angolo (*a sinistra*). I risultati non erano influenzati dal fatto che alcune stimolazioni (*cerchietti*) fossero orientate in modi che i soggetti non avevano visto negli esperimenti precedenti: nel processo immaginato gli oggetti non saltano da una posizione appresa all'altra, ma passano per posizioni intermedie proprio come gli oggetti che ruotano nel mondo fisico. Quando i ricercatori non hanno deliberatamente rispettato la corrispondenza tra la stimolazione e l'orientazione che il soggetto stava immaginando (*a destra*), i tempi di risposta sono aumentati con la deviazione angolare. Può darsi che, per valutare la forma, i soggetti dovessero immaginare una rotazione di compensazione.

La distorsione delle forme normali ha permesso di verificare la fedeltà con cui vengono conservate le caratteristiche di uno stimolo fisico sottoposto a rotazione mentale. Dopo avere indicato ai soggetti in quale direzione dovevano immaginare di ruotare una delle forme normali, i ricercatori presentavano nella nuova orientazione una forma normale o una sua distorsione. I soggetti erano in grado di individuare anche le minime variazioni, e questo fa pensare che nel processo di rotazione mentale venga conservata molta della ricchezza strutturale della forma originaria.

quale versione si trattasse. Dato che possedevamo già un'estesa messe di dati sulla velocità di rotazione di ciascun soggetto, abbiamo potuto ideare prove tagliate su misura dei singoli soggetti. Dagli esperimenti precedenti ricavammo a quali tempi e angoli ciascun soggetto sarebbe stato più pronto a reagire a una forma. In metà delle prove, che chiamammo «a stimolazione attesa» (*probe-expected trial*), la forma venne presentata secondo un'angolazione e un tempo che corrispondevano esattamente alla rotazione mentale in corso; nell'altra metà venne invece mostrata secondo un'angolazione che differiva da quella che, in base ai calcoli, doveva essere l'orientazione immaginata in quel momento. Questo secondo tipo di prove è stato chiamato «a stimolazione inattesa» (*probe-unexpected trial*).

Se un soggetto, durante lo svolgimento della rotazione mentale, necessariamente immagina un oggetto in orientazioni intermedie, dovrebbe rispondere sempre con la stessa velocità a una stimolazione correttamente sincronizzata e orientata nello spazio, qualunque sia l'angolo preso in considerazione. I tempi di reazione registrati nelle prove a stimolazione attesa erano coerenti con la nostra ipotesi: i tempi di risposta nelle prove in cui la forma veniva presentata con una orientazione attesa erano virtualmente costanti e tutti intorno al mezzo secondo per qualunque angolo.

Le prove a stimolazione attesa presentano un'altra caratteristica istruttiva. Metà delle stimolazioni erano ruotate, rispetto alla forma normale, di multipli di 60 gradi, come era già stato fatto nel precedente esperimento con un singolo stimolo. Le altre stimolazioni «attese» venivano mostrate con orientazioni insolite, cioè ruotate di multipli dispari di 30 gradi. Se, come riteniamo, la rotazione mentale non salta in modo discontinuo da un angolo all'altro, ma procede con continuità attraverso stati corrispondenti ad angoli intermedi, i tempi di risposta a stimolazioni opportunamente sincronizzate, ma orientate secondo angoli insoliti si sarebbero dovuti approssimare a quelli per angoli ai quali i soggetti erano abituati. I tempi di risposta quasi identici nei casi di orientazioni familiari e in quelli di orientazioni insolite hanno confermato l'ipotesi.

Quando la stimolazione veniva invece orientata in un modo diverso da quello atteso, i tempi di risposta crescevano linearmente con la deviazione angolare. Evidentemente i soggetti dovevano immaginare una ulteriore rotazione correttiva quando veniva loro presentata una forma che non corrispondeva alla loro rappresentazione mentale in quell'istante; la correzione richiedeva un tempo proporzionale alla entità della deviazione angolare. Questa scoperta fornisce una ulteriore prova a sostegno del fatto che è la corrispondenza tra orientazione immaginata e orientazione mostrata a essere importante al fine di ottenere tempi di risposta brevi e costanti, e non la familiarità che i soggetti hanno con le forme in tutte le loro orientazioni. I tempi occorrenti alla correzione della rotazione stanno a indicare che, benché i soggetti fossero allenati a quel tipo di esperimento, non potevano valutare le forme mostrate loro senza prima attuare una rotazione mentale.

In complesso, i nostri risultati equivalgono a una prova oggettiva dell'esistenza di un processo mentale che riproduce la rotazione di oggetti nel mondo fisico. Le due scoperte fondamentali sono la relazione lineare fra il tempo di reazione e la differenza di orientazione, quando si confronta la forma intrinseca di due stimoli, e l'uniforme rapidità di risposta quando un oggetto viene presentato in una posizione che corrisponda a quella immaginata nella rotazione mentale. Abbiamo raccolto dati precisi e riproducibili su un fenomeno apparentemente soggettivo, del tipo di quelli che, in passato, venivano considerati al di fuori del campo di ricerca della psicologia sperimentale.

Rimangono interrogativi sulla natura delle trasformazioni mentali che abbiamo studiato. Sebbene abbiamo stabilito che, per determinare l'identità di due oggetti che vengono mostrati con orientazioni diverse, può essere necessario immaginare una rotazione che passa attraverso stadi intermedi, non intendiamo con questo sostenere che la rotazione sia continua in senso rigorosamente matematico, cioè che passi attraverso tutti i possibili angoli intermedi. Non si conosce, infatti, la base neurofisiologica delle immagini mentali e della loro manipolazione interna.

Rimangono anche altri interrogativi. Quando le immagini mentali vengono trasformate, quanti particolari conservano dei corrispondenti oggetti fisici? Secondo esperimenti recenti le rappresentazioni mentali possono conservare molto della ricchezza strutturale dei loro corrispettivi materiali. In una ricerca di Cooper e Peter Podgorny (allievo di Cooper) i soggetti erano in grado di distinguere le forme ruotate da quelle normali non solo quando le stimolazioni differivano per riflessioni nello spazio, ma anche quando la distinzione era questione di sottili e imprevedibili distorsioni (*si veda l'illustrazione in basso nella pagina a fronte*).

A parte qualche questione irrisolta, la stretta corrispondenza che abbiamo trovato fra le rotazioni mentali e i loro corrispettivi nel mondo fisico conduce inevitabilmente a meditare sulle funzioni e sull'origine dell'immaginazione spaziale dell'uomo. Potrebbe non essere prematuro avanzare l'ipotesi che l'immaginazione spaziale si sia sviluppata come riflesso della fisica e della geometria del mondo esterno. Forse, nel corso della storia evolutiva, le regole che governano le strutture e i movimenti del mondo fisico sono state incorporate nel meccanismo percettivo umano, producendo corrispondenze dimostrabili tra le immagini mentali e i loro corrispettivi fisici. Ecco che iniziamo a scoprire un meccanismo mentale altrettanto preciso ed elegante dello schematismo innato che Chomsky ha posto a fondamento del linguaggio.

I grandi vascelli spagnoli

Anche se il primo posto fra le più grandi navi da guerra dell'epoca della vela viene assegnato in generale agli inglesi, era il vascello spagnolo «Santísima Trinidad» la nave da guerra più grande e temibile del tempo

di John D. Harbron

Il 21 ottobre 1805, nell'ultima e più cruenta battaglia navale dell'epoca velica, la flotta britannica, diretta dall'ammiraglio Lord Horatio Nelson, annientò al largo di Capo Trafalgar le forze unite delle flotte francese e spagnola. L'esito di tale battaglia segnò la fine della Spagna come potenza marittima e impose prepotentemente gli inglesi alla considerazione di tutto il mondo facendoli collocare al primo posto per l'architettura navale, la progettazione di navi e la capacità di manovra del XVIII secolo. Eppure per i marinai e i capitani di vascello del periodo la nave da guerra più grandiosa e temibile presente a Trafalgar in quel fatale pomeriggio era la *Santísima Trinidad*, una piattaforma galleggiante di 1900 tonnellate di dislocamento, irta di cannoni, con fiancate in mogano spesse 60 centimetri e con un equipaggio di circa 1200 fra marinai e soldati di marina. La *Santísima Trinidad* fu elogiata in più di un'occasione dallo stesso Nelson e fu oggetto di ammirazione da parte di molte altre flotte nemiche della Spagna per la sua architettura e la sua robustezza in mare. Una valutazione del suo successo come strumento da guerra può essere desunta dalla sorte riservatale dagli inglesi dopo essersene impadroniti nel corso della battaglia. Nel timore che potesse essere ripresa dagli spagnoli, che pure erano fuggiti con quanto restava della loro flotta, e incapaci di rimorchiarla nelle condizioni disastrate in cui si trovava in conseguenza dei danni che le erano stati inflitti dalla battaglia e da una violenta tempesta, gli inglesi la affondarono due giorni dopo aprendole delle falle nello scafo. La conoscenza che abbiamo di questa splendida nave si fonda perciò inevitabilmente su ricostruzioni tratte da testimonianze del tempo.

La storia ha trascurato la costruzione navale spagnola a causa delle sconfitte subite dagli iberici nel XVIII secolo. Le sconfitte erano però dovute a una migliore tattica navale e a una maggiore disciplina dei marinai inglesi, e non a una peggiore progettazione delle navi spagnole. In una tale valutazione storica non obiettiva si ignora il successo della Spagna nell'istituire e mantenere rotte navali mercantili per il Nuovo Mondo.

I navigatori spagnoli della fine del Quattrocento e dei primi decenni del Cinquecento, imbevuti dallo zelo religioso della Controriforma, fecero progressi sorprendenti nell'esplorazione del mondo. Inoltre, già nel 1600, ossia a soli ottant'anni dallo sbarco di Hernán Cortés in Messico, esistevano nell'America spagnola da 40 a 50 comunità di coloni. La maggior parte degli insediamenti poteva essere rifornita e difesa solo dal mare; il loro successo è attribuibile almeno in parte a una grandissima tradizione di costruzioni navali che durò per quasi trecento anni.

La colonizzazione del Nuovo Mondo non avrebbe potuto aver luogo con l'aiuto della nave a cui viene assegnato di solito un indebito rilievo nelle esposizioni storiche della costruzione navale spagnola: la *galera*, o galea, una nave dal fondo piatto, a propulsione sia remica sia velica, che dominò la guerra navale nel Mediterraneo per tutto il Medioevo. (L'ultima grande battaglia dell'epoca del remo, in cui una flotta mista formata da galee della Spagna e di altri paesi europei sconfisse una flotta turca, fu combattuta al largo della costa greca, a Lepanto, nel 1571.) La galea non era certo adatta per le lunghe e spesso difficili traversate dell'Atlantico. Colombo e gli esploratori che seguirono il suo esempio usarono diverse varietà della cosiddetta nave rotonda, o cocca, e verso la fine del Cinquecento questo nuovo tipo aveva sostituito completamente la galea della flotta spagnola. Lo scafo molto più grande della nave a vari ponti consentiva il trasporto di più provviste, più uomini, più cannoni e una velatura più abbondante, tutte cose di cui vi era necessità nei lunghi viaggi mercantili e di scoperta. Inoltre poiché lo scafo era arrotondato, appiattito al fondo e inclinato verso l'interno lungo i ponti superiori, questa nave teneva molto meglio il mare.

La nave rotonda aveva tre alberi e il progetto di base venne realizzato in due versioni specializzate: il *galeón* e il *navío*. I *galeónes*, o galeoni spagnoli, erano navi mercantili che trasportavano grandi carichi d'oro, argento e provviste, oltre a un numero consistente di uomini d'equipaggio e di passeggeri. I *navíos* erano navi da guerra, con scafi di grande spessore ed equipaggiate con cannoni; i *navíos* più potenti e duraturi erano i vascelli, che occupavano una posizione avanzata in ogni scontro. (Di qui l'espressione usata un tempo di «nave di linea» come sinonimo di «nave da battaglia»; dalla corrispondente espressione inglese *«line-of-battle ship»* deriva la parola inglese moderna *«battleship»*, da cui anche l'espressione italiana «nave da battaglia».) Il grande vascello spagnolo della metà e della fine del Settecento rappresentò il culmine di secoli di costruzione della cocca a tre alberi. La *Santísima Trinidad* era l'esempio più bello di *navío*.

Com'era costruito il vascello spagnolo del Settecento? Le tecniche adottate dagli spagnoli erano simili a quelle usate dalle altre potenze navali europee. La spina dorsale dell'ossatura del vascello era costituita dalla chiglia e dal paramezzale, una struttura di sostegno costruita al di sopra della chiglia per rafforzare l'intelaiatura circostante. Il paramezzale terminava in corrispondenza del dritto di prora nella parte anteriore della nave e del dritto di poppa in quella posteriore. Le coste dell'ossatura della nave erano fissate alla chiglia e l'intera struttura era bloccata con vari tipi di fasciame, tanto all'interno quanto all'esterno dello scafo.

Il fasciame comprendeva i corsi di tavola del fasciame esterno, che erano grandi elementi longitudinali trasversali che sostenevano l'ossatura interna, e i bagli di ponte trasversali, che collegavano solidamente le coste su un fianco della nave con quelle sull'altro fianco.

La struttura era tenuta assieme da caviglie di quercia o di mogano e da bulloni di ferro pudellato, tutti prodotti sul posto nei cantieri navali reali. In una nave da guerra di grandi dimensioni come la *Santísima Trinidad* i bulloni di ferro erano lunghi fino a 180 centimetri. Essi venivano fatti passare attraverso fori praticati

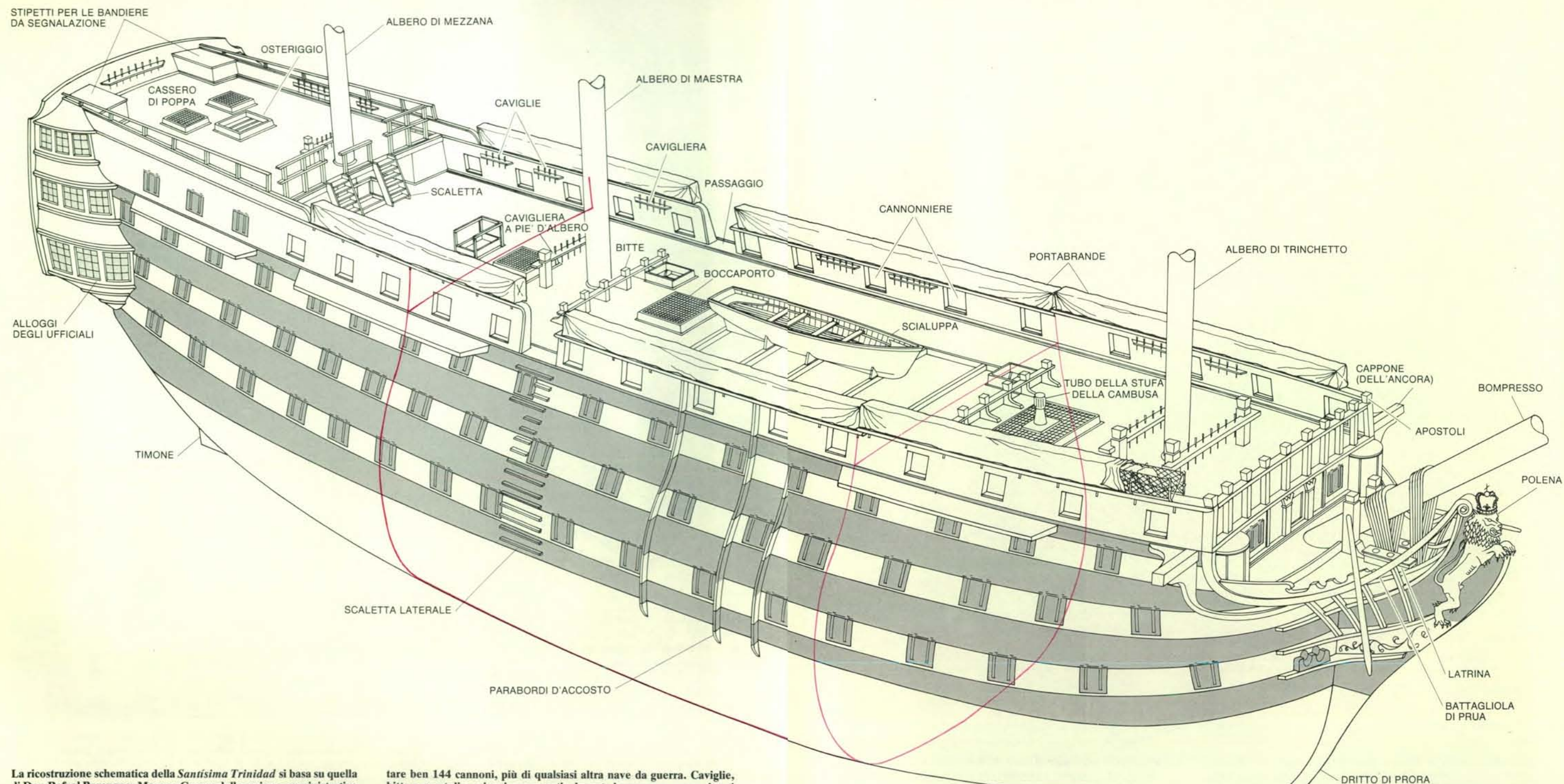
La *Santísima Trinidad*, raffigurata con le vele spiegate in un dipinto coevo di Don Alejo Berlinguero de la Marca y Gallego, era la nave più grande fra quelle che parteciparono alla battaglia di Trafalgar svoltasi il 21 ottobre 1805. Fu costruita nel cantiere navale spagnolo dell'Avana nel 1769, con legno di mogano e di pino del Messico e dell'America Centrale. Le finestre visibili nella poppa sono quelle degli alloggi degli ufficiali, che erano arredati elegantemente, ma poco protetti. La tattica navale del Settecento richiedeva spesso che una nave da guerra che attaccava manovrasse in modo tale che la prima salva dell'artiglieria potesse colpire la poppa della nave nemica. Il quarto ponte dotato di cannoni che si vede nel dipinto fu aggiunto durante un raddobbo posteriore, che fece della *Santísima Trinidad* la nave da guerra più armata del Settecento. La nave portava in battaglia anche un equipaggio composto da 1200 fra marinai e soldati di marina. Pur avendo subito pesanti bordate da parte delle artiglierie navali inglesi, non fu mai affondata in battaglia. Il dipinto si trova al Museo Naval di Madrid.

nel legno ed erano bloccati da dadi di ferro pudellato o di legno. Mentre la costruzione procedeva in bacino di carenaggio, le vele venivano preparate cucendo assieme varie strisce di tela di lino, e con la canapa si intrecciava il sartiame per sostenere gli alberi e assicurare le vele.

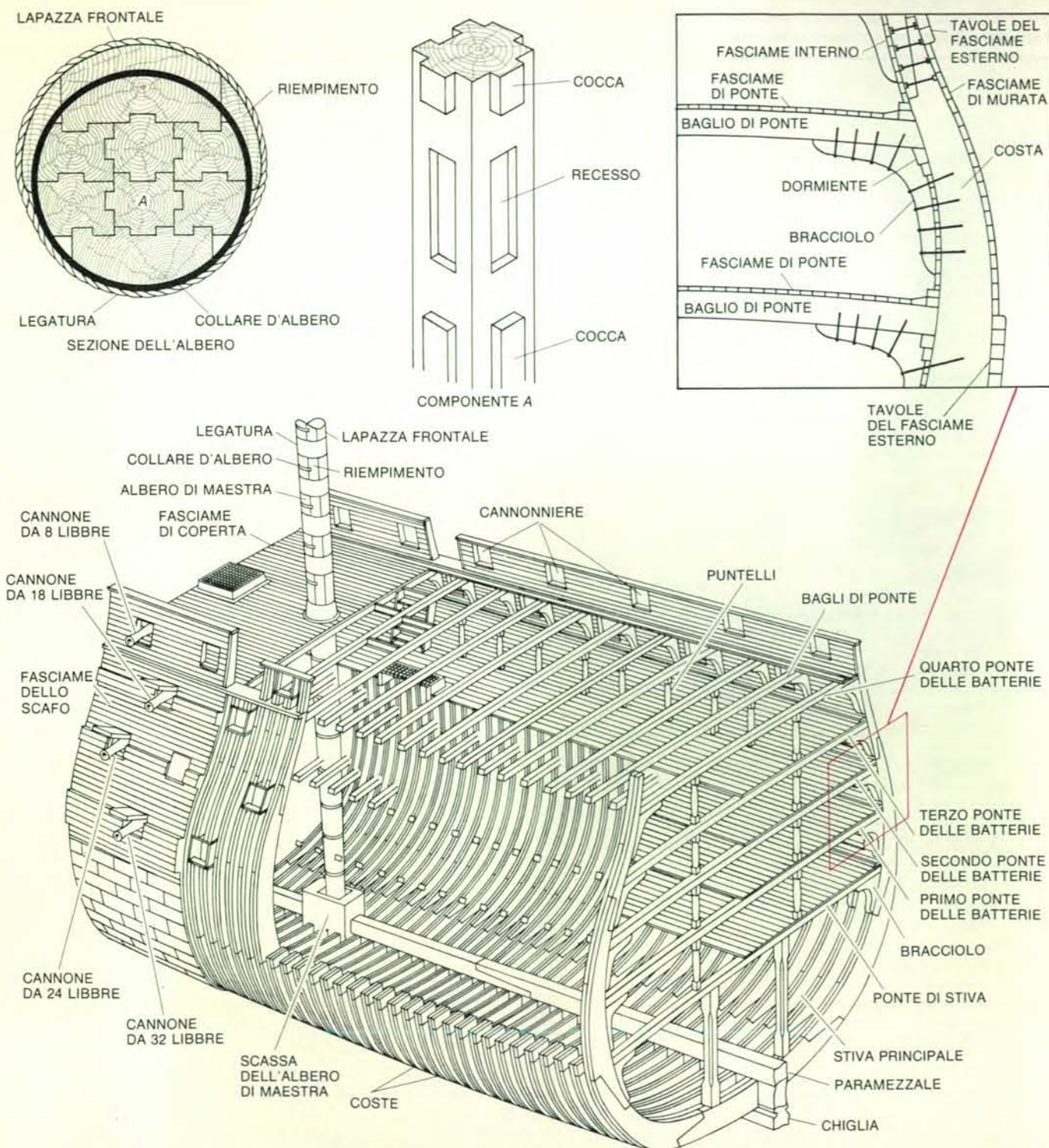
La cosa più importante per il marinaio era il modo in cui l'intera struttura - lo scafo di legno, le vele e il sartiame - avrebbe funzionato in mare. L'effetto delle sollecitazioni continuamente mutevoli del vento e delle onde su migliaia di pezzi di legno bloccati da caviglie di legno e da bulloni di ferro faceva del *navío*, molto più che delle navi a vela di oggi, un sistema in un equilibrio dinamico piuttosto precario, che doveva essere tenuto costantemente sotto controllo e sottoposto a continue correzioni da parte del capitano e dell'equipaggio. Il vascello, percorrendo gli oceani, cigolava, gemeva, si sollevava e ricadeva sulla superficie del mare, in un modo che non viene più sperimentato dai marittimi attuali. Persino coloro che viaggiano sulle alte navi di oggi hanno il grande vantaggio della relativa rigidità di uno scafo d'acciaio.

La chiave del successo del *navío* spagnolo risiedeva nei materiali di qualità superiore di cui i costruttori di navi spagnoli disponevano. Lo scafo delle navi da guerra francesi e inglesi del tempo era fatto di legno di quercia, e gli alberi e i pennoni erano di legno di pino. L'impero ispano-americano riforniva invece l'*Armada española* di legni duri più robusti, come il mogano, che cresceva a Cuba e nelle foreste costiere dell'attuale Honduras. Il mogano è di gran lunga più resistente della quercia al marciume secco, provocato da un fungo che consuma la cellulosa del legno stagionato lasciandone una struttura fragile che si sbriciola facilmente. Tutte le navi in legno sono soggette a tale grave inconveniente, cosicché per una marina da guerra era essenziale poter contare su una fonte affidabile di legno duro, non solo per la costruzione di nuove navi ma anche per raddobbare le vecchie.

La pronta disponibilità di legno duro in grado di dar buona prova di sé in mare per periodi di tempo relativamente lunghi fra un raddobbo e l'altro era perciò un deciso vantaggio per una potenza navale. In un periodo in cui tanto gli inglesi quanto i francesi stavano cominciando a dubitare seriamente della loro capacità di produrre o procurarsi legno di quercia e di pino in quantità sufficiente per la costruzione e la manutenzione delle navi, la Spagna poté continuare a contare su abbondanti forniture di legni duri dal suo impero americano. Gran parte del legname veniva trasportato nel cantiere reale dell'Avana, che costruì 74 dei 221 *navíos* a tre alberi varati dalla Spagna nel corso del Settecento e che, fra il 1770 e il 1780, diventò il più grande cantiere del mondo per la costruzione di vascelli. La *Santísima Trinidad*, il cui scafo e i cui ponti furono costruiti per intero con mogano di Cuba e dell'Honduras, fu varata all'Avana nel 1769.

La quantità di legname che veniva usata nella costruzione delle navi da guerra era enorme. Più di 3000 alberi, ciascuno dei quali forniva 180 metri di tavole, erano necessari per costruire la cosiddetta nave da guerra di terzo rango, il cavallo da tiro delle flotte europee. I pini per gli alberi e i pennoni (le antenne orizzontali che sostenevano le vele quadre), crescevano in abbondanza nell'attuale

La ricostruzione schematica della *Santísima Trinidad* si basa su quella di Don Rafael Berenguer Moreno Guerra della sezione amministrativa della marina spagnola (che attinse a sua volta a descrizioni contemporanee). La nave era lunga 62,3 metri e pesava circa 1900 tonnellate; dopo l'aggiunta del quarto ponte delle batterie, avrebbe potuto trasportare ben 144 cannoni, più di qualsiasi altra nave da guerra. Caviglie, bitte e apostoli a cui assicurare sartie durante le manovre erano situati in punti strategici in coperta. Le brande non usate dall'equipaggio venivano riposte di giorno nelle impavesate lungo le murate del ponte di coperta. Le linee in colore indicano i limiti dello spaccato di pagina 94.

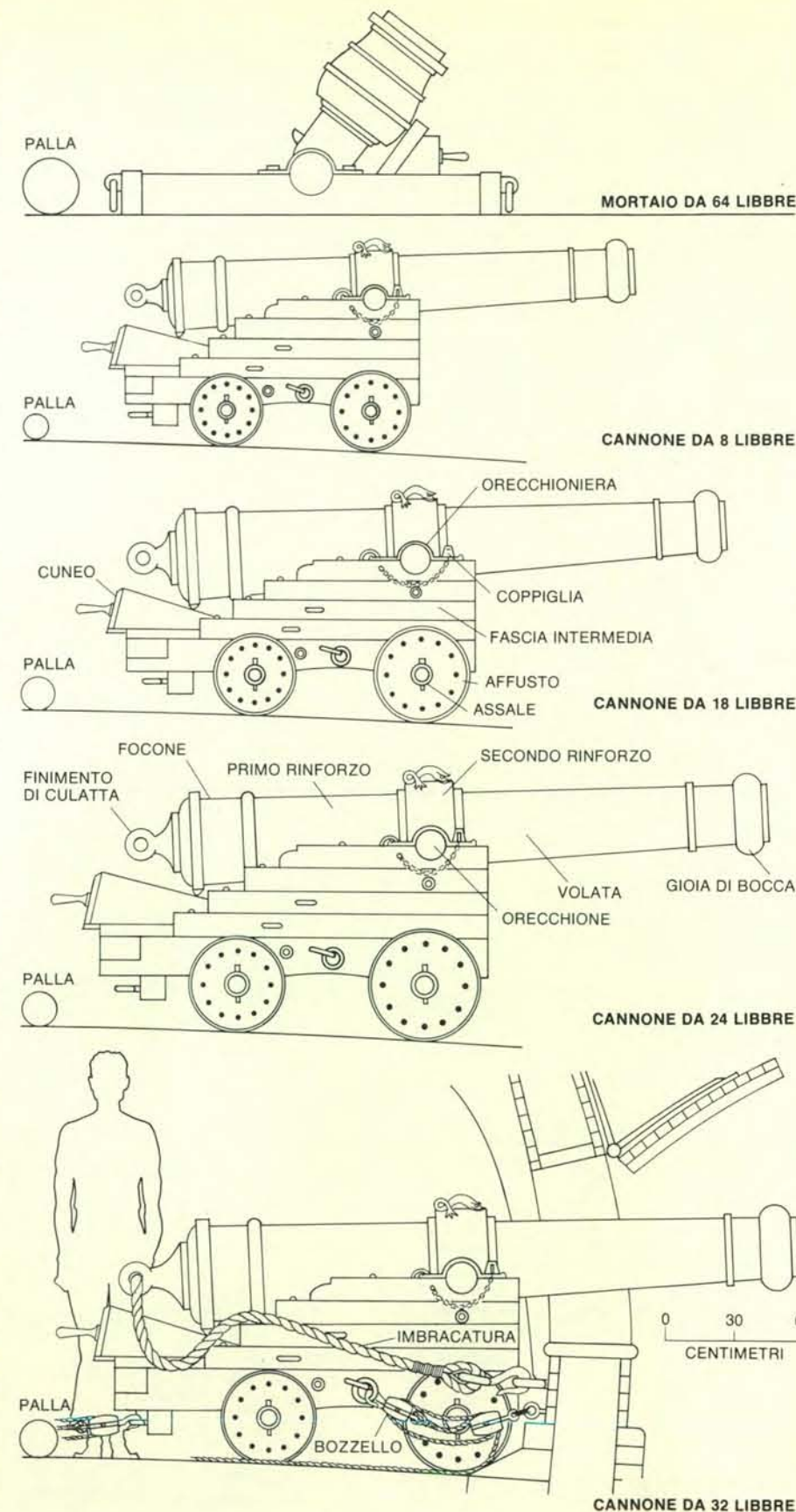
Messico. Ben 40 pini erano necessari per produrre 22 pennoni di una nave da guerra a tre alberi di terzo rango. Il cantiere navale era il più grande ente manifatturiero dell'età preindustriale alla fine del XVIII secolo e non può esserci dubbio sul fatto che le grandi navi da guerra erano i «manufatti» più grandi prodotti regolarmente prima della Rivoluzione Industriale. Nondimeno, le 60 navi a tre ponti e tre alberi che combatterono a Trafalgar non pesavano complessivamente più di 120 000 tonnellate, ossia pressappoco la stazza lorda di una superpetroliera moderna.

Quando si rendeva necessario un raddobbo, esso non si limitava ai ponti, alle cabine, agli alberi e alle vele, come nelle navi e nelle imbarcazioni di oggi, dallo scafo in acciaio, plastica, fibra di vetro o alluminio. La manutenzione delle navi da guerra del Settecento richiedeva la sostituzione periodica dell'intelaiatura fondamentale dello scafo; di fatto, le navi da

Lo spaccato della *Santísima Trinidad* mostra l'intelaiatura e la disposizione dei ponti a mezza nave, esposti lungo i piani indicati dalle linee in colore nell'illustrazione delle due pagine precedenti. La costruzione venne iniziata a partire dalla chiglia e dal paramezzale; le coste poi aggiunte e bloccate da puntelli, bagli di ponte e corsi di tavole del fasciame esterno. La maggior parte delle scorte per un viaggio venivano riposte nella stiva principale. Il ponte di stiva, o falso ponte, ospitava le attrezzature del chirurgo e un deposito di munizioni in cui erano conservate cariche già predisposte per i cannoni. I cannoni più grandi e pesanti erano montati nel primo ponte immediatamente al di sopra del ponte di stiva. Gli alloggi per l'equipaggio erano soprattutto sui ponti di mezzo e superiore; gli uomini appendevano le brande fra i cannoni.

guerra che combatterono a Trafalgar non erano le stesse di quando erano state costruite e varate per la prima volta. Il metodo usato sarebbe familiare a un attuale proprietario di yacht. Delle pulegge, sulla riva, venivano fissate ai cavi appropriati del sartiame della nave e, rimossi i cannoni, il *navío* veniva abbattuto in carena, ossia inclinato su un fianco. Poteva allora incominciare il lavoro di sostituzione delle tavole marce; quando la sostituzione non era necessaria, ci si limitava a scrostarle di tutti gli organismi marini che erano cresciuti su di esse e a verniciarle con varie sostanze conservanti. Il fondo della nave veniva spesso rivestito con fodere di rame per fornirgli un'ulteriore protezione contro il marciume oltre che contro le brume o teredini, molluschi bivalvi del genere *Teredo* che scavano gallerie nelle parti immerse degli scafi di legno.

Questi lavori di manutenzione assicuravano ovviamente alla nave una vita molto più lunga: la *Santísima Trinidad* aveva 36 anni quando fu affondata a Trafalgar, e la nave ammiraglia di Nelson, la *Victory*, ne aveva 40. Il fondo della nave pulito, foderato di rame e non ingombro da organismi marini, vegetali e animali, rendeva la nave più veloce e più facilmente manovrabile in azione. Una maggiore velocità non era però sempre un fatto positivo. Nella battaglia di Trafalgar la nave britannica *Royal Sovereign*, appena foderata di rame, sopravanzò di molto due linee della flotta inglese e si trovò a vedersela da sola con il vascello spagnolo *Santa Ana*, armato di 112 cannoni, ricevendone bordate micidiali.

Il costo della manutenzione delle navi aveva una grande incidenza sull'economia del tempo di guerra degli stati marittimi europei. Per esempio, lo studioso di storia economica José P. Merino Navarro, della Fondazione universitaria spagnola di Madrid, ha stimato che la *Victory*, la cui costruzione era costata agli inglesi nel 1765 circa 63 000 sterline, costò loro di manutenzione sino alla fine delle guerre napoleoniche, nel 1815, circa 372 000 sterline. Nondimeno, per proteggere i suoi interessi nelle Americhe e nelle Filippine spagnole, la Spagna mantenne in tutto il mondo più centri di carenaggio e di manutenzione di quanti non ne avessero francesi o inglesi, e le navi spagnole erano di solito in condizioni molto migliori rispetto alle loro controparti inglesi. L'ammiraglio inglese Lord Cuthbert Collingwood, uno fra i più grandi capitani marittimi del tempo, inviò più di una volta a Nelson e all'Ammiragliato dispacci sul pessimo stato delle navi inglesi dell'epoca. Alcuni studiosi di storia navale hanno scritto che gli inglesi non avrebbero potuto combattere e vincere un'altra battaglia come quella di Trafalgar senza prima aver combattuto il marciume del legno. Pare evidente che un tale compito avrebbe messo a dura prova le risorse finanziarie della nazione.

Il rinnovamento della marina da guerra spagnola nel XVIII secolo si inquadrò in un ammodernamento di tutte le istitu-

Cannoni e mortai di cui erano dotate le navi da guerra del Settecento. Sul primo ponte delle batterie, il più basso, la *Santísima Trinidad* portava 30 cannoni da 32 libbre (circa 14,5 chilogrammi), i più grandi del servizio navale alla fine del Settecento; erano lunghi tre metri e potevano a volte colpire un bersaglio alla distanza di due chilometri e mezzo. Il ponte di mezzo delle batterie della *Santísima Trinidad* ospitava 32 cannoni da 24 libbre, e il ponte superiore aveva un armamento misto composto da 32 cannoni da 18 libbre e 26 da otto, oltre a mortai. Quando fu aggiunto il quarto ponte, il numero dei cannoni a bordo fu portato da 120 a 144. Le imbracature per posizionare i cannoni sono esemplificate schematicamente per il cannone da 32 libbre.

zioni spagnole realizzato dal re Carlo III, uno dei «despoti illuminati» del secolo. Il sovrano introdusse metodi migliori di organizzazione e di pianificazione a lungo termine dell'intera industria spagnola, compresi i cantieri navali reali. Uno fra i miglioramenti più efficaci fu la standardizzazione dei progetti di navi da guerra. All'inizio del Settecento ogni nave era la creazione unica del suo progettista. Fra il 1770 e il 1780 i *navíos* erano raggruppati invece in sei ranghi principali, o classi, e le navi all'interno di ciascun rango erano costruite secondo disegni architettonici comuni. Inoltre i progetti per navi di ranghi diversi richiedevano scafi, strutture dei ponti armati di cannoni e forma delle vele simili. Ne risultò in modo evidente un'accresciuta efficienza del cantiere navale, in un'epoca in cui la Spagna aveva bisogno di una flotta in rapida espansione di nuove navi per difendere un immenso impero nel suo lungo conflitto con l'Inghilterra.

Ciascuno dei progetti era una soluzione dell'inevitabile compromesso fra il bisogno di una piattaforma galleggiante stabile per le batterie di cannoni e la capacità di manovrare rapidamente e di dare alla nave un rapido abbrivo con la propulsione velica. La classificazione per ranghi si fondava primariamente sul numero dei ponti armati di cannoni e sul numero di cannoni a bordo. I vascelli di primo rango, come la *Santísima Trinidad*, erano le navi da guerra più grandi e avevano a bordo le

La suddivisione delle navi in ranghi si fondava in gran parte sul numero di ponti con cannoni e sul numero di cannoni. I vascelli di primo rango erano le navi da guerra più potenti; per lo più portavano da 80 a 110 cannoni distribuiti su tre ponti. Le navi del secondo rango avevano da 80 a 98 cannoni su tre ponti e quelle del terzo da 74 a 80 cannoni su due ponti. Solo le navi di primo, secondo e terzo rango venivano chiamate navi di linea. Le navi del quarto, quinto e sesto rango erano più leggere e di conseguenza più veloci, e venivano usate con compiti di crociera ed esplorazione e come corrieri. Le navi di quarto rango avevano da 50 a 60 cannoni sui due ponti, quelle di quinto da 32 a 44 su un singolo ponte e quelle di sesto non più di 20-28 cannoni su un singolo ponte.

artiglierie più potenti. La *Santísima Trinidad* poteva trasportare fino a 144 cannoni ed era l'unica nave da guerra in servizio con quattro ponti armati di cannoni. Tutti gli altri vascelli di primo rango avevano tre ponti armati di cannoni e tre alberi. Le navi di secondo rango avevano da 80 a 98 cannoni distribuiti su tre ponti, e quelle di terzo rango avevano da 74 a 80 cannoni distribuiti su due ponti. Le navi appartenenti ai primi tre ranghi erano quelle designate in generale come «navi di linea».

Il quarto, quinto e sesto rango erano assegnati abitualmente a squadre di navi più piccole, fra cui le fregate, adibite a compiti di crociera e di esplorazione; trasportando meno cannoni, esse potevano raggiungere velocità più elevate. Le navi di quarto rango avevano fra 50 e 60 cannoni distribuiti su due ponti, quelle di quinto ne avevano fra 32 e 44 su un solo ponte e quelle di sesto fra 20 e 28 su un solo ponte. Le navi del sesto rango erano le più veloci della marina spagnola, e spesso veniva loro affidato il compito di portare la posta reale e le istruzioni reali fra la Spagna e le sue lontane colonie.

Dalla chiglia al ponte di coperta, un *navío* spagnolo era un alto e largo edificio a cinque piani. Sopra la sentina, nella tipica nave di primo rango del tempo, vi era la stiva principale, e al di sopra di essa il falso ponte, o ponte di stiva. A questo seguivano i tre ponti principali per le batterie di cannoni. Il grande peso dei cannoni causava spesso squilibri che accentuavano il rollio con il mare mosso; in tali circostanze il più basso fra i ponti con batterie, specialmente in navi inglesi del tempo, finiva spesso sotto la linea di galleggiamento.

Il ponte di stiva, immediatamente al di sopra della sentina, era la parte più sicura della nave; perciò vi si trovavano tutte le attrezzature del chirurgo per la cura delle ferite e per eseguire le molte amputazioni immediate che si rendevano necessarie durante le cruente battaglie navali. Sulle navi inglesi il ponte di stiva era di solito dipinto di rosso per celare i fiumi di sangue che scorrevano dal tavolo operatorio. Nella *Santísima Trinidad* sulla paratia al di sopra del tavolo operatorio era appeso un grande crocifisso, simbolo della fede religiosa alla quale più d'uno offriva un braccio, una gamba o la vita. Il ponte di stiva ospitava anche un deposito di munizioni in cui si trovavano i cartocci, cariche di polvere premisurate e pronte all'uso per i cannoni.

Per contrastare la tendenza della nave a rollare, lo scafo del *navío* spagnolo era più largo all'altezza del primo ponte armato di cannoni che non all'altezza di quello medio e di quello superiore. Il ponte armato più basso era perciò il più stabile e ospitava i cannoni più grandi presenti sulla nave, pezzi lunghi tre metri che sparavano palle di 32 libbre, ossia di circa 14,5 chilogrammi. (La libbra spagnola era di 460 grammi, contro i 453,59 della libbra inglese.) La *Santísima Trinidad* portava 30 di questi cannoni, 15 su ciascun fianco. A breve distanza i suoi cannoni potevano forare uno scafo di quercia dello spessore di 90 centimetri e potevano causare ancora più danno a una nave nemica quando la carica veniva raddoppiata o triplicata. Le palle da 32 libbre potevano a volte colpire un bersaglio a una distanza di due chilometri e mezzo.

Oltre ai cannoni, il ponte medio e superiore comprendevano anche la maggior parte dei miseri alloggi assegnati all'equipaggio. Le brande venivano sospese fra i cannoni e dovevano essere riposte prima della battaglia. Una cambusa era situata al centro della nave e ivi una singola stufa era disponibile all'equipaggio per cuocervi il cibo e far bollire l'acqua. I *navíos* spagnoli erano anche eccessivamente affollati: la tattica navale spagnola richiedeva un gran numero di uomini, cosa che aveva come conseguenza una mescolanza spesso ingovernabile di veri uomini di mare e di marinai improvvisati. Alla battaglia di Trafalgar la *Santísima Trinidad* aveva un equipaggio in eccesso, 1200 uomini fra marinai e soldati di marina, molti dei quali ancora convalescenti dall'epidemia di malaria e di colera che aveva flagellato la Spagna meridionale fra il 1802 e il 1804. Di contro la nave inglese di primo rango *Victory*, che era a sua volta sovraffollata, aveva a bordo solo 900 uomini, e tutti marinai di mestiere.

Gli alloggi degli ufficiali a bordo del *navío* erano molto lussuosi, riflettendo la rigorosa struttura gerarchica della società spagnola. Gli ufficiali avevano loro cabine private e vivevano e pranzavano quasi con le stesse comodità cui erano abituati a terra. La tavola nella cabina di poppa era imbandita con buoni cibi e buon vino. Belle vetrine contenevano porcellane, caraffe e carte nautiche. Tutti questi oggetti di lusso andavano però spesso in frantumi nel corso di una battaglia, tanto più in conseguenza dell'abitudine degli inglesi di manovrare le navi per sparare salve contro la poppa di un *navío*.

Un equipaggio numeroso era necessario anche quando non si combatteva poiché quasi tutto il lavoro di governo della nave veniva eseguito a mano. All'inizio di un viaggio, tutto ciò che era a bordo, comprese le provviste della nave, le munizioni e persino i cannoni più pesanti, doveva essere sistemato a mano. Una volta salpata l'ancora, i marinai erano impegnati costantemente sulle sartie per serrare e spiegare le decine e decine di vele. Le pompe montate sul ponte di mezzo delle batterie per mantenere la sentina a un livello di sicurezza venivano azionate a mano e, durante la battaglia, altre pompe venivano mantenute in funzione per spegnere gli incendi e per lavare i ponti, spesso inondati di sangue. Le pulegge a bordo alleviavano un po' il carico di lavoro, e per calare e salpare le ancore e per orientare pesanti vele si usavano verricelli; anch'essi, però, erano azionati a mano.

A differenza degli inglesi, spagnoli e francesi consideravano le navi da guerra piattaforme da combattimento per soldati di marina, oltre che come vere e proprie piattaforme di artiglieria. Questa strategia, come s'è accennato, ebbe spesso come conseguenza una diffusa indisciplina tra l'equipaggio, la quale probabilmente svolse un ruolo decisivo nella sconfitta degli spagnoli a Trafalgar. In linea di principio, la tattica dei capitani spagnoli consisteva nel far fuoco nel sartiame, per abbattere gli alberi delle navi nemiche. Ridotte all'incapacità di manovra, queste dovevano essere abbordate e conquistate in un combattimento corpo a corpo.

La tattica inglese consisteva invece nello sparare direttamente allo scafo della nave nemica. Anche in conseguenza del maggiore rollio delle loro navi rispetto a quelle spagnole, gli inglesi erano più pronti a sostituire la tradizionale miccia con l'acciarino per accendere le polveri dei cannoni. L'acciarino accendeva la carica quasi istantaneamente dando all'artigliere una maggiore probabilità che il proiettile partisse anticipando il nuovo movimento del rollio della nave.

Da entrambe le parti i piani di battaglia venivano fatti ben comprendendo che la potenza di fuoco della nave da guerra del Settecento era particolarmente efficace quando il bersaglio era alla portata dei tiri con l'alzo a zero. Quando un *navío* riusciva a manovrare in modo da trovarsi a poppa di una nave nemica, una singola salva di cannoni con proiettili da 32 libbre cancellava la poppa ben decorata ma poco protetta. Un tale tiro di sbarramento poteva inoltre spazzar via gli artiglieri dal ponte della nave nemica, spostare i pesanti cannoni e disseminare all'intorno migliaia di schegge di legno, ciascuna delle quali poteva essere un'arma letale.

Quando due navi da guerra si trovavano a fianco in battaglia, le bocche dei cannoni venivano ritirate attraverso le cannoniere e si sparava dall'interno della nave. Il risultato era devastante. I ponti delle batterie, alti appena 180 centimetri, comprimevano il rumore degli spari in un rombo che causava agli artiglieri una sordità permanente. Palle di cannone roventi dovevano essere caricate nella bocca dei cannoni con utensili a mano, e i cannoni, il cui calore costituiva anch'esso un grave rischio per l'incolumità degli uomini dopo alcune salve, dovevano essere riportati a mano nella posizione di fuoco dopo ogni rinculo. I ponti erano invasi dal fumo dei cannoni e spesso accadeva che qualche membro dell'equipaggio venisse travolto mortalmente in conseguenza del rinculo di qualche cannone che non aveva visto attraverso il fumo denso.

Nonostante la terribile potenza di fuoco, i *navíos* stessi erano in grado di resistere a ore di battaglia a breve distanza. La *Santísima Trinidad*, per esempio, svolse il compito di nave ammiraglia spagnola durante il blocco di Gibilterra nel 1779-1782 e combatté al largo del Capo di San Vincenzo, in Portogallo, nel 1797, oltre che a Trafalgar. Bordate con doppia o tripla carica sparate dai vascelli inglesi non riuscirono ad affondarla e anche se gli inglesi riuscirono infine a prevalere, la maggior potenza degli spagnoli fece una strage di molti loro vascelli.

Sarebbe ingiusto dare un giudizio storico definitivo sul successo del *navío* spagnolo senza menzionarne il ruolo nell'esplorazione e nella scoperta. Vascelli spagnoli circumnavigarono il globo in viaggi scientifici commissionati da Carlo III che ebbero un'utilità non minore dei viaggi molto più conosciuti dei capitani inglesi James Cook e George Vancouver. Le spedizioni spagnole diedero l'avvio a studi di biologia marina, di oceanografia e di tecnica della navigazione. I capitani di marina, così come i botanici, i geografi e i cartografi che viaggiavano a bordo dei *navíos*, prepararono minuziose relazioni sulle loro osservazioni.

I capitani di marina spagnoli non furono semplicemente strateghi militari. Due fra i diciotto capitani in servizio a Trafalgar furono fra i migliori navigatori, esploratori, cartografi e matematici del loro tempo. Dionísio Alcalá Galiano fu un innovatore il cui nome è commemorato in molti nomi di località lungo la costa occidentale del Cile e del Canada. Il *brigadier* (grado della marina spagnola intermedio fra capitano di vascello e capo di squadra) Don Cosmé Damián Churruca, un basco che riscrisse molti testi di navigazione della marina spagnola, esplorò la costa occidentale dell'America Meridionale e diede importanti contributi alla matematica. Come Lord Nelson, loro contemporaneo e avversario inglese, entrambi rimasero uccisi a bordo delle loro navi nel sanguinoso pomeriggio di Trafalgar.

(RI)CREAZIONI AL CALCOLATORE

di A. K. Dewdney

Squali e altri pesci combattono una guerra ecologica sul pianeta Wa-Tor

Da qualche parte, in una direzione che si può solo chiamare ricreativa e a una distanza limitata soltanto dal proprio valore programmatorio, il pianeta Wa-Tor nuota tra le stelle. È un pianeta a forma di toro, cioè di ciambella, ed è interamente coperto dall'acqua. Gli abitanti principali di Wa-Tor sono squali e altri pesci, così chiamati perché queste sono le creature della Terra a cui maggiormente assomigliano. Gli squali di Wa-Tor mangiano i pesci più piccoli, di cui sembra vi sia sempre grande abbondanza.

Questo semplice ecosistema potrebbe apparire stabile, quasi soporifico, se non fosse per il fatto che le popolazioni degli squali e degli altri pesci subiscono violente oscillazioni. Molte volte, nel passato, la popolazione di questi ultimi è stata quasi interamente divorata, altre volte gli squali hanno sfiorato l'estinzione (anche quando gli altri pesci abbondavano). Eppure, sia gli squali sia gli altri pesci sopravvivono. Per capire perché, ho scritto un programma che simula le loro attività alimentari e riproduttive.

Prima di mostrare questi ritmi ecologici sullo schermo, però (*si veda la figura della pagina a fronte*), ho meditato a lungo sulle regole e i particolari del programma WATOR. Un giorno, a pranzo, mi trovai a chiacchierare con David Wiseman, il mago dei sistemi di calcolo del mio dipartimento all'Università dell'Ontario occidentale. Dopo avergli descritto il progetto, notai che Magi (così viene chiamato Wiseman) sorrideva in modo enigmatico. Il mattino seguente, tutto fiero, mi fece entrare nel suo ufficio per mostrarmi un programma già pronto a funzionare.

«Guarda», disse, e premette un tasto. Un assortimento inizialmente casuale di squali e altri pesci guizzava da un punto all'altro in modo apparentemente caotico. Alcuni squali non riuscivano a nutrirsi e scomparivano; altri avevano una prole vorace quanto loro. Alcuni pesci più piccoli, abbastanza fortunati da occupare una regione in cui in quel momento non c'erano squali, si moltiplicavano formando un grande banco. A questo punto, un gran numero di squali scoprivano il banco, si raggruppavano ai suoi lati e vi penetravano per un poco facendosi strada mangiando. Pochi minuti dopo, la situazione sullo schermo di Magi era la seguente: 578 pesci piccoli e solamente 68 squali.

Qualcuno entrò nell'ufficio di Magi e corse subito fuori. Nemmeno cinque minuti dopo, la stanza era piena di gente che faceva il tifo per gli squali. A poco a poco, i piccoli pesci furono circondati da un muro di squali, mentre in un altro punto dello schermo un banco poco numeroso di piccoli pesci si moltiplicava senza essere notato. Mormorii si levarono quando infine questi pesci scomparvero e gli squali, agitandosi alla ricerca di preda, incominciarono a morire uno a uno. Mi venne l'idea di modificare le regole per consentire agli squali di mangiarsi l'un l'altro, ma mi resi conto che un'alimentazione parossistica non avrebbe prolungato in modo significativo la loro esistenza e avrebbe potuto mettere a repentaglio la vita dell'altro piccolo banco. Quando infine due squali, nel loro vagare, si imbatterono in esso, il ciclo ricominciò.

Il programma di Wa-Tor non è né molto lungo né difficile da scrivere. I lettori che possiedano un calcolatore personale, anche con poca esperienza di programmazione, si sentiranno gratificati quando infine il codice sarà stato scritto, corretto e fatto girare. Parametri quali i tempi di alimentazione, i periodi di carestia e la dimensione delle popolazioni iniziali possono essere stabiliti in principio; poi, non c'è che da mettersi comodi e osservare come un miscuglio inizialmente disorganizzato di squali e pesci più piccoli assuma a poco a poco configurazioni ecologiche.

Nel programma WATOR compaiono molte regole semplici che governano il comportamento sia degli squali sia degli altri pesci. Queste creature nuotano in un oceano a griglia rettangolare i cui lati opposti sono identificati a due a due. Questo significa semplicemente che se uno squalo o un altro pesce occupa un punto della griglia a destra e decide di nuotare verso est (cioè verso destra), riapparirà nel corrispondente punto della griglia a sinistra. Lo spazio bidimensionale senza confini che ne risulta è un toro, l'effettiva superficie di Wa-Tor (*si veda l'illustrazione di pagina 100*). Per scrivere il proprio programma WATOR si può scegliere una qualsiasi opportuna dimensione per la griglia oceano. Per esempio, Magi, il cui programma gira su un calcolatore VAX, ha definito un oceano largo 80 punti e alto 23. La mia versione di WATOR, scritta per un IBM PC, usa un più modesto oceano 32 per 14.

Il tempo trascorre con salti discreti, che chiamerò crononi. Durante ogni cronone uno squalo o un altro pesce può spostarsi in direzione nord, est, sud o ovest verso un punto adiacente, sempre che il punto non sia già occupato da un membro della sua stessa specie. La scelta è effettuata da un generatore di numeri casuali. Per un pesce non predatore la scelta è semplice: scegliere a caso un punto libero adiacente e spostarcisi. Se tutti i quattro punti adiacenti sono occupati, il pesce non si muove. Dato che predare i pesci ha la priorità sul semplice movimento, le regole per uno squalo sono più complesse: scegliere a caso uno dei punti adiacenti occupati dalle eventuali prede, spostarcisi e divorarle. In loro mancanza, lo squalo si sposta esattamente come un pesce non predatore, evitando gli altri squali.

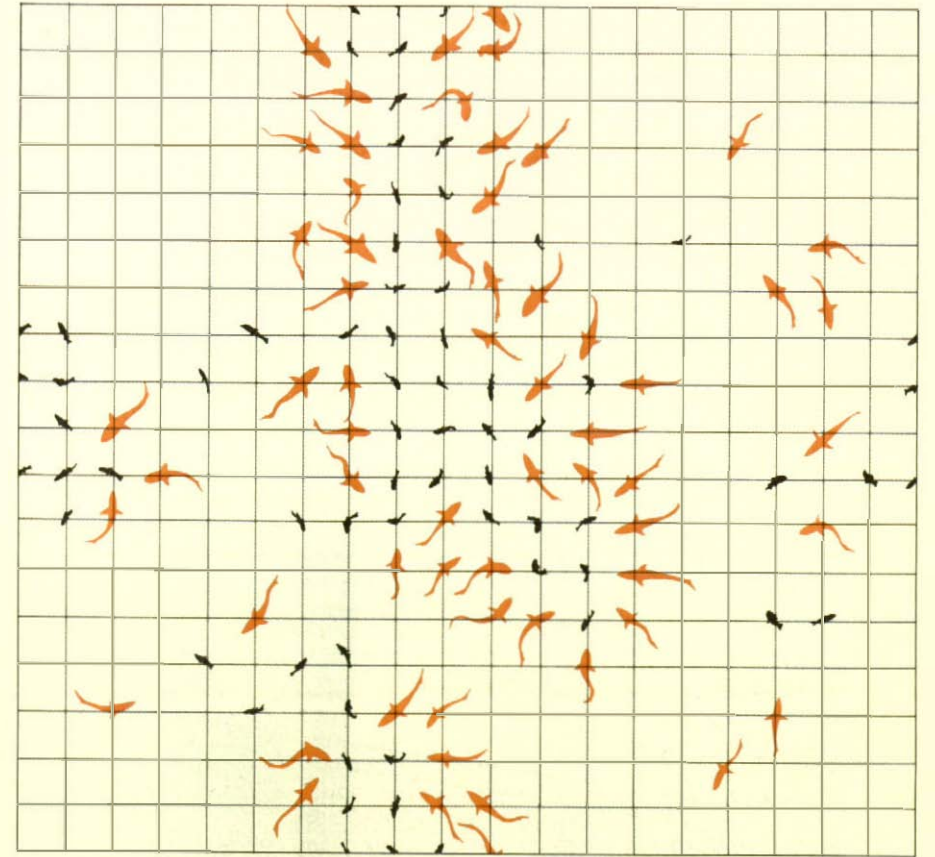
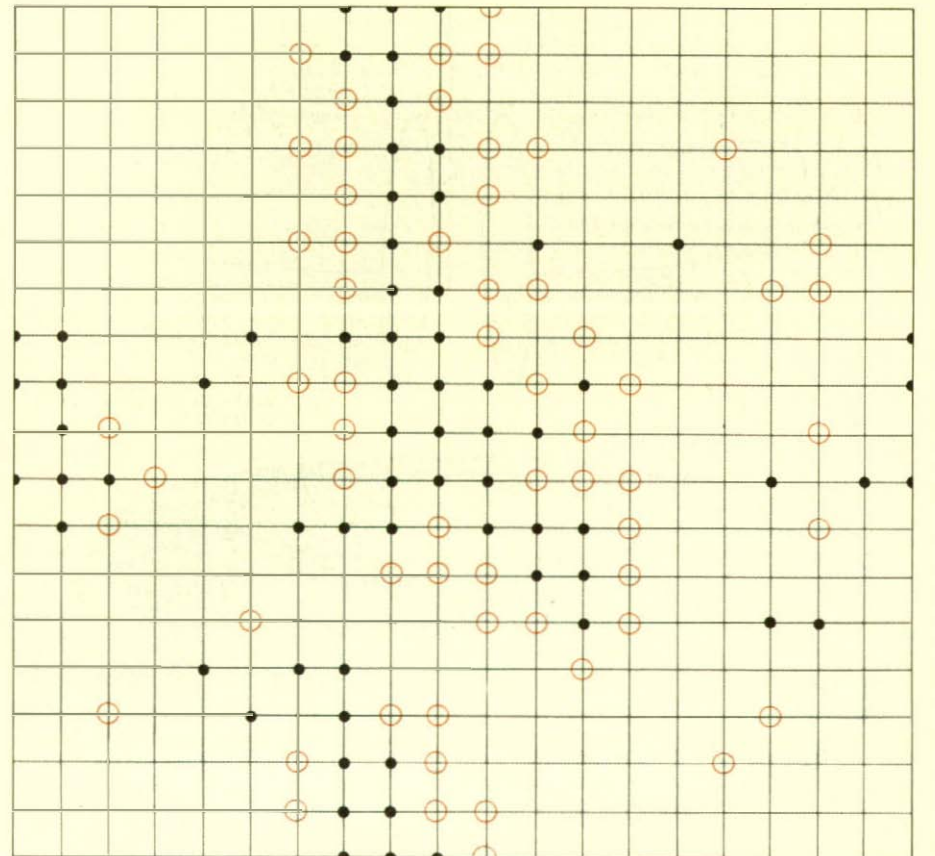
Il creatore di WATOR sceglie 5 parametri per avviare una data simulazione. I parametri *npesci* e *nsquali* rappresentano il numero di pesci (possibili prede degli squali) e di squali all'inizio dell'esecuzione. Il programma distribuisce i pesci e gli squali a caso e in modo più o meno uniforme sulla superficie del pianeta. I parametri *pprole* e *sprole* designano il numero di crononi che devono passare prima che, rispettivamente, un pesce e uno squalo abbiano un figlio. (Entrambe le specie sono evidentemente partenogeniche.) Infine *inedia* indica il numero di crononi che uno squalo ha a disposizione per trovare di che cibarsi. Se lo squalo non riesce a mangiare prima che sia trascorso questo tempo, muore e scompare dalla circolazione. Durante ogni cronone, WATOR muove una volta ciascun pesce e ciascuno squalo e mostra i risultati sullo schermo.

Io e Magi abbiamo assistito a un gran numero di sceneggiature a cinque parametri in cui l'oceano di Wa-Tor si sovrappopolava di pesci finché gli squali si moltiplicavano al punto da mangiare tutti gli altri pesci e poi morire. In altre occasioni abbiamo visto divorati tutti i pesci di un grande banco; gli squali che si erano ingozzati finivano col morire di fame, senza scoprire mai un gruppetto di pesci lì vicino. In qualche caso abbiamo visto la relazione preda-predatore conservarsi per due o anche tre cicli prima del definitivo tracollo nella popolazione degli squali. Nulla, però, nei parametri scelti per quelle sceneggiature dava qualche suggerimento sulle caratteristiche che garantirebbero un ecosistema eterno. Come avevano fatto a sopravvivere gli abitanti di Wa-Tor?

Qualcuno ha detto che biologia vuol dire destino. Magi e io siamo tentati di dichiarare che ecologia vuol dire geometria, almeno per quanto riguarda il pianeta Wa-Tor. Il definitivo destino di una data sceneggiatura non sembra dipendere dalla distribuzione casuale iniziale di un certo numero di squali e di altri pesci. E neppure sembra dipendere in modo accidentale dall'effettivo movimento casuale degli squali e di quei pesci. La probabilità del tracollo di una popolazione sembra invece seguire strettamente la geometria pesce-squalo che si manifesta sui nostri schermi: più cresce l'organizzazione e la localizzazione delle due popolazioni, più è probabile la definitiva condanna dell'ecosistema. Meditando su questo argomento, ci siamo chiesti come avremmo potuto scegliere i cinque parametri in modo da rompere la geometria. Ci venne allora un lampo di intuizione: nel caso gli squali si fossero radunati ai lati di un banco di pesci, un modo per spezzare la geometria risultante sarebbe potuto consistere nel diminuire la frequenza di riproduzione degli squali. L'assembramento stesso, dopo tutto, derivava più dalla riproduzione che dal movimento.

Prima di formulare questa ipotesi, avevamo scelto tempi di riproduzione grosso modo uguali per gli squali e gli altri pesci. L'equilibrio nei ritmi riproduttivi, pensavamo, avrebbe dato come conseguenza un equilibrio nelle popolazioni. Questo genere di ragionamento vago è probabilmente all'origine di molte calamità dell'attuale mondo tecnologico. In ogni caso, io misi 20 squali e 200 altri pesci nel mio oceano 32 per 14 e stabilii che i pesci si riproducessero ogni tre crononi, mentre gli squali si riproducevano solo dopo 10 crononi. Il tempo di morte per inedia degli squali venne fissato, in modo più o meno arbitrario, in tre crononi. Fummo ricompensati, dopo aver osservato per 15 minuti il mio programma, piuttosto lento, dal vedere che la popolazione si ricostituiva dopo il declino iniziale. Per di più, la geometria, pur essendo ancor presente, era più suggerita che definita. I banchi erano conglomerati informi con margini sfilacciati e sia gli squali sia gli altri pesci si aggiravano casualmente in alcuni punti dello schermo.

Lasciai girare il programma tutto il pomeriggio, dando qualche occhiata di tanto in tanto mentre mi occupavo di questioni più importanti. Il programma girò tutta la notte e quando entrai nel mio ufficio, dopo la lezione del mattino, trovai che sia gli squali sia gli altri pesci continuavano la loro esistenza ciclica. Questa volta era proprio Wa-Tor!

Ci sono molti modi per implementare un programma WATOR, ma forse il più semplice consiste nell'utilizzare matrici bidimensionali. Io uso cinque matrici a cui possiamo dare il nome di PESCI, SQUALI, MOSSAPESCE, MOSSASQUALO e INEDIA. Queste matrici, tutte di 32 per 14, registrano le posizioni e le età degli squali e degli altri pesci. PESCI (I, J) rappresenta la presenza o l'assenza di un pesce nel punto con coordinate (I, J). Se non c'è un pesce, la posizione ha il valore −1. In caso contrario, contiene una registrazione dell'età in crononi del pesce che vi si trova. Lo stesso schema viene usato nella matrice SQUALI per registrare le posizioni e le età degli squali. La matrice MOSSAPESCE registra, per ogni posizione, se un pesce è stato spostato in quella posizione durante il ciclo computazionale in corso. Questa registrazione impedisce al programma di muovere un pesce due volte durante lo stesso cronone. MOSSASQUALO adempie la stessa fun-

Un'immagine realistica di squali che predano pesci

Un'immagine più facilmente programmabile, in cui i cerchi rappresentano gli squali e i punti i pesci che essi predano

zione per gli squali. La matrice chiamata INEDIA registra il momento in cui uno squalo ha mangiato per l'ultima volta. Se in una posizione non ci sono squali, il valore corrispondente è −1.

Il modo più semplice per visualizzare ciò che avviene su Wa-Tor è una linea di caratteri sullo schermo per ogni riga delle matrici; un vuoto in una posizione significa che non è occupato. Un punto (.)

Il pianeta toroidale Wa-Tor e la sua rappresentazione su una carta piana (o sullo schermo piano di un calcolatore)

rappresenta un pesce e uno zero (0) rappresenta uno squalo. Pur essendo apparentemente limitata, questa visualizzazione è sorprendentemente ricca di informazione e piacevole da guardare.

Nella fase iniziale di WATOR, i pesci, eventuali prede degli squali, e gli squali nel numero richiesto sono distribuiti uniformemente nell'oceano toroidale. Il programma, poi, attraversa ciclicamente i tre segmenti o sottoprogrammi descritti qui di seguito; ogni ciclo del programma dura un cronone di tempo.

I pesci nuotano e si riproducono:
Per ogni pesce della matrice PESCI, il programma elenca le posizioni adiacenti non occupate e sposta il pesce su una di esse a caso. PESCI allora deve essere posta a −1 nella vecchia posizione e all'attuale età del pesce nella nuova posizione. La matrice MOSSAPESCE viene aggiornata nel modo descritto prima. Se l'età del pesce è uguale a *pprole*, il programma mette un nuovo pesce nella vecchia posizione e dà età 0 a entrambi i pesci. MOSSAPESCE registra il nuovo pesce. Se tutte le posizioni adiacenti sono occupate, il pesce non si muove né si riproduce.

Gli squali predano e si riproducono:
Per ogni squalo della matrice SQUALI, il programma elenca le posizioni dei pesci adiacenti (se ce ne sono). Lo squalo sceglie a caso una di queste posizioni, vi si sposta e mangia il pesce. Questo non significa solo che il programma deve modificare SQUALI e MOSSASQUALO come ha modificato PESCI e MOSSAPESCE, ma anche che deve assegnare −1 alla corrispondente posizione nella matrice PESCI. Inoltre, a INEDIA in quella posizione è assegnato il valore 0. Se non ci sono pesci da predare, lo squalo si muove come uno di questi pesci. Se l'età dello squalo è uguale a *sprole*, nasce un nuovo squalo esattamente come nasce un nuovo pesce.

Visualizzazione:
Il programma esamina la matrice PESCI e la matrice SQUALI, poi visualizza un punto per ogni pesce e un cerchio per ogni squalo. La visualizzazione può essere effettuata tutta insieme oppure può essere divisa in due parti: una eseguita dopo che si sono spostati i pesci, l'altra dopo che si sono spostati gli squali.

Per popolare l'oceano iniziale, il programmatore costruisce un ciclo che genera due numeri casuali per *npesci* volte. I numeri sono in proporzione alle dimensioni orizzontale e verticale dell'oceano che si vuole avere. In ognuna delle posizioni casuali così scelte, il programma mette un pesce nella matrice PESCI e gli assegna un'età compresa tra 0 e *pprole*. La distribuzione degli squali avviene in modo analogo. In entrambi i casi, si verifica se la posizione è già occupata. Dando età casuali sia agli squali sia agli altri pesci, si ottiene che essi si riproducano in momenti casuali, con un effetto di maggior naturalezza. Senza questa precauzione, si assisterebbe a un improvviso raddoppio del numero degli squali e dei pesci, cosa sconcertante e innaturale.

Forse i programmatori alle prime armi possono trovare la precedente descrizione un po' troppo generale per farsi un'idea chiara di come scrivere un programma WATOR. Possono iniziare allora a scrivere quello che si può chiamare programma dell'ubriaco barcollante. Un programma di questo genere potrebbe essere formato da un unico ciclo (per esempio un ciclo *while*) con sette istruzioni. Le istruzioni qui sono scritte in un generico linguaggio algoritmico. Gli assegnamenti sono indicati da una freccia rivolta verso sinistra e le variabili X e Y sono le coordinate di un ubriaco barcollante. Esse variano a seconda dell'intero casuale assegnato a una variabile *direzione*. A seconda che questo intero sia uguale a 0, 1, 2 o 3, l'ubriaco (un punto sullo schermo) si muove verso nord, est, sud o ovest.

direzione ← parte intera
di (*random* × 4)
se *direzione* = 0 allora $X \leftarrow X + 1$
se *direzione* = 1 allora $X \leftarrow X - 1$
se *direzione* = 2 allora $Y \leftarrow Y + 1$
se *direzione* = 3 allora $Y \leftarrow Y - 1$
visualizzazione (X, Y)

Se il vostro generatore di numeri casuali produce un numero decimale *casuale* compreso tra 0 e 1, questo algoritmo lo ridurrà a un numero decimale di valore compreso tra 0 e 0,3999. La parte intera del numero deve essere 0, 1, 2 o 3.

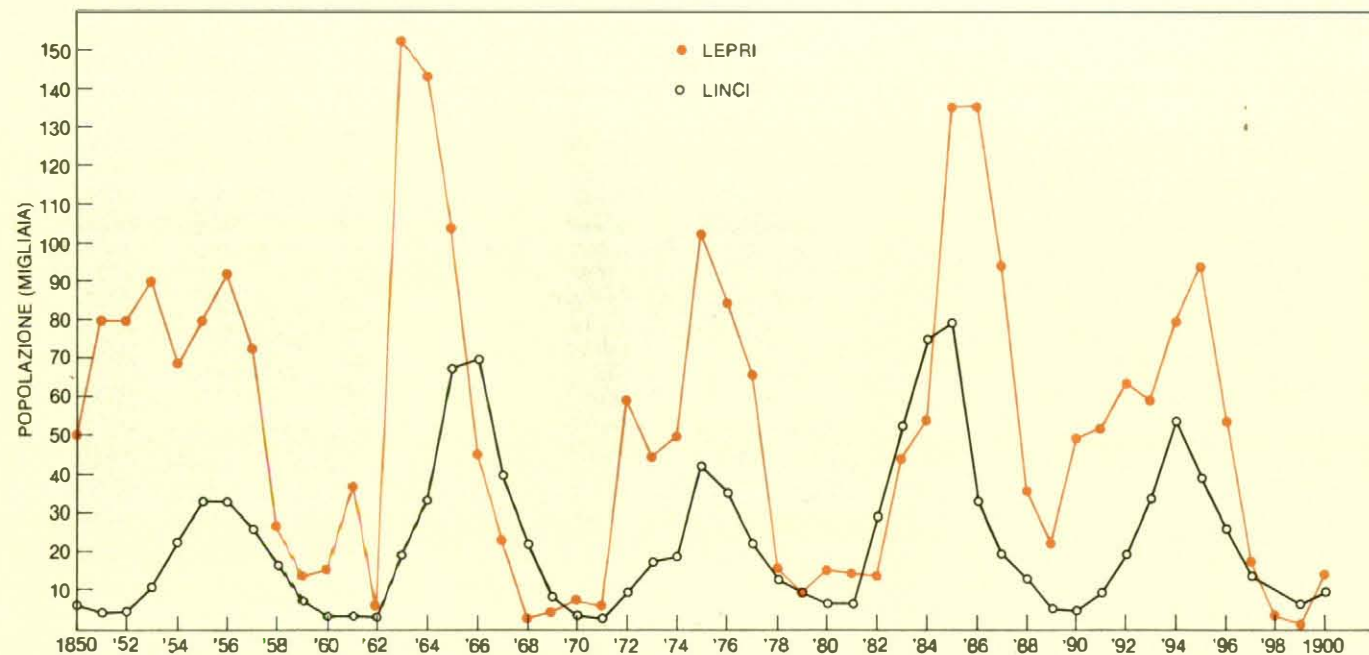
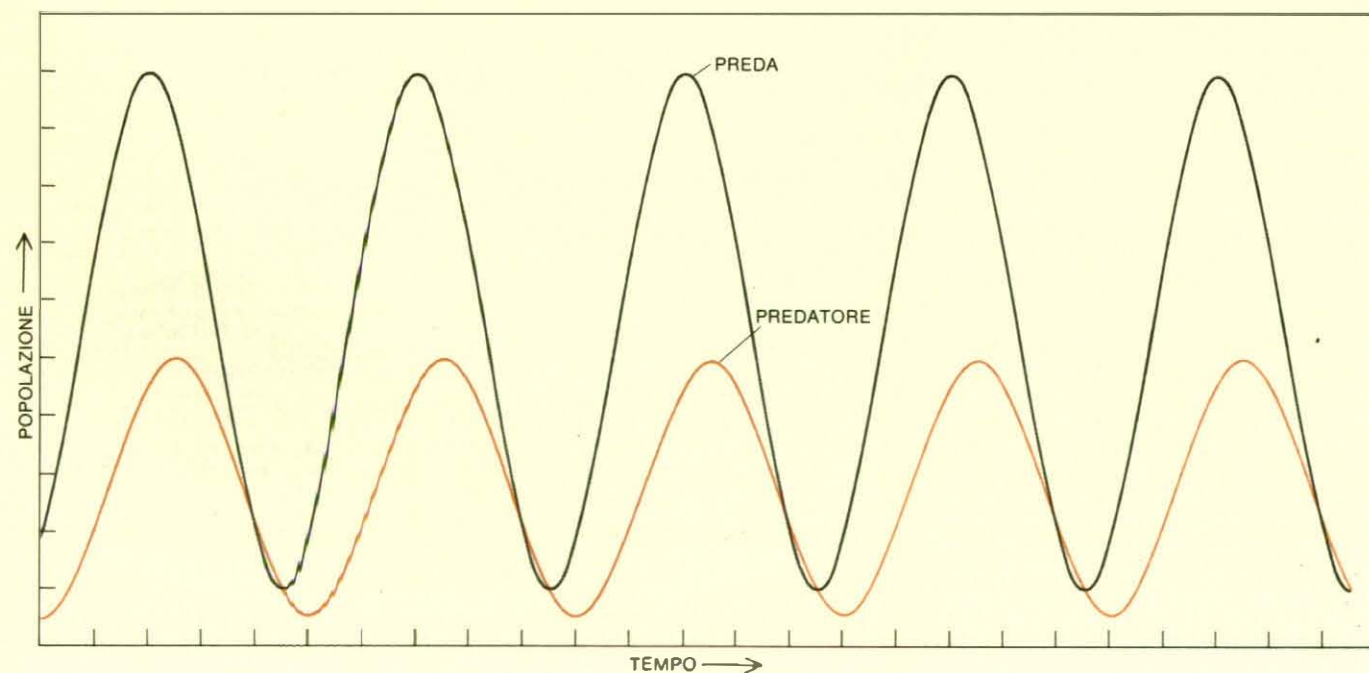
Non posso certo affermare che stare a guardare un punto luminoso che vaga sullo schermo sia paragonabile al dramma ecologico degli squali e dei pesci che essi predano, ma scrivendo questo programma si può avere un'idea di come si potrebbero costruire le varie parti di WATOR.

I programmatori esperti che leggano questo articolo penseranno ad altri modi per scrivere il programma WATOR. La quantità di elaborazione può essere ridotta di molto usando liste concatenate per seguire gli squali e i pesci che essi predano.

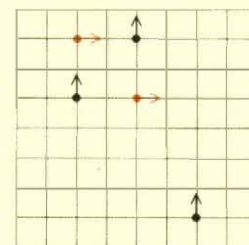
WATOR può fornire qualche indicazione sulle popolazioni animali che vivono sulla Terra. Sappiamo che le piccole popolazioni hanno un'alta probabilità di estinzione e, anche se né i predatori né la preda scompaiono, quasi certamente essi sottostanno a cicliche variazioni nel numero. In semplici ecosistemi predatore-preda, le popolazioni dei predatori e delle prede seguono due cicli sovrapposti di massimi e di minimi di popolazione. Le dimensioni delle popolazioni di lepri e di linci registrate dalla Compagnia della Baia di Hudson dal 1847 al 1903 nel territorio subartico canadese seguono questo schema (*si veda l'illustrazione della pagina seguente*). Le cifre indicano il numero di animali delle due specie catturati di anno in anno. È presumibile che questi numeri siano proporzionali alle effettive dimensioni delle due popolazioni durante questo periodo. Se lo sono, i cicli si spiegano facilmente come il risultato delle predazioni operate dalle linci in una popolazio-

ne sempre crescente di lepri, popolazione che comincia a declinare con l'aumento del numero di linci. Ben presto comincia a esservi meno cibo per le linci, che iniziano a mo
meno (o tutt'e due le cose). Quando il numero delle linci si riduce, le lepri iniziano di nuovo a moltiplicarsi.

Messo a confronto con questo diagramma è un insieme di curve che rappresentano una soluzione alle equazioni di Lotke-Volterra. Queste equazioni sono state formulate per la prima volta nel 1931 dal matematico Vito Volterra. Esse presuppongono quello che si potrebbe chiamare un predatore continuo, cioè continuamente a caccia di una preda continua. Le soluzioni di queste equazioni mostrano una variazione ciclica che, a prima vista, sembra riprodurre i dati empirici relativi alle lepri e alle linci. I biologi, però, non concordano sul fatto che l'entità delle lepri e delle linci sia spiegabile con un ragionamento così semplice. In primo luogo, sono in gioco almeno altri due nemici delle lepri: i microbi e l'uomo.

È del tutto sensato, comunque, compilare statistiche degli squali e degli altri pesci di Wa-Tor, e io e Magi l'abbiamo fatto. I nostri grafici più recenti relativi alle popolazioni degli squali e dei pesci tendono ad assomigliare ai diagrammi delle linci e delle lepri più di quanto facciano le soluzioni di Lotke-Volterra. Eppure, continua a lasciarci perplessi l'instabilità mostrata sul lungo periodo da talune combinazioni di parametri. Forse qualche lettore, lavorando sul suo programma WATOR, potrà dare qualche ulteriore illuminazione. Esiste qualche regola generale che permetta di prevedere, data una combinazione di parametri, se l'ecosistema risultante sarà stabile? In che misura le fluttuazioni cicliche seguono le equazioni di Lotke-Volterra?

Numero di linci e di lepri (in migliaia) catturate dalla Compagnia della Baia di Hudson dal 1850 al 1900

Una relazione teorica predatore-preda: una soluzione alle equazioni di Lotke-Volterra

L'oceano di Wa-Tor è toroidale per una ragione semplicissima: è molto più semplice scrivere un programma per un oceano che non abbia limiti o coste. Se l'oceano deve essere largo, per esempio, 32 unità, è facile usare numeri modulo 32 come coordinate X dei pesci e degli squali. Se questi hanno come coordinata X 31 e appaiono sul margine destro dello schermo durante un cronone, possono anche avere come coordinata X 32 = 0 e apparire sul margine sinistro nel cronone successivo. Lo stesso per l'altra coordinata.

L'oceano toroidale di Wa-Tor dà luogo ad alcuni effetti molto strani. Il primo ha a che fare con un difetto di una delle prime versioni del mio programma WATOR. Questo difetto faceva sì che ogni pesce nuotasse di un'unità verso nord e ogni squalo nuotasse di un'unità verso est durante ogni cronone di tempo. Uno squalo, quindi, arrivava a catturare un pesce solo se si trovava a occupare la stessa posizione della sua preda. Dato l'oceano qui sotto, quanti pesci non venivano mai catturati da uno squalo?

Un altro rompicapo prevede squali e pesci intelligenti. Supponiamo che a ogni mossa ogni squalo e ogni pesce si spostino su uno qualsiasi dei loro quattro punti adiacenti. Ne risulta che un singolo pesce, se è abbastanza intelligente, può sempre sfuggire a un singolo squalo, per quanto intelligente sia il predatore. Nell'oceano toroidale di Wa-Tor, due squali a caccia di un pesce isolato possono produrre due risultati differenti. Se dotate ogni creatura di tutta l'intelligenza che volete, arrivando anche a permettere che gli squali cooperino nella caccia, riuscite a scoprire un modo in cui il pesce possa cavarsela? Il risultato non dipende dalle dimensioni dell'oceano.

A proposito dei perceptroni (si veda il fascicolo di novembre 1984), alcuni lettori hanno trovato applicazioni e altri si sono sentiti stimolati ad approfondire l'argomento. E. Manning di Stanford, Connecticut, ha costruito 10 anni fa un particolare perceptrone, per convertire immagini reali nei quadrati digitalizzati della retina di un perceptrone. Manning è stato uno dei pochi che hanno notato un errore nel perceptrone a finestra con rettangoli multipli della figura in alto di pagina 178 del numero di novembre 1984: le ultime quattro configurazioni del demone dovrebbero essere metà in colore e metà bianche.

Gary D. Stormo, un ricercatore del Dipartimento di biologia molecolare, cellulare e dello sviluppo dell'Università del Colorado a Boulder, ha usato il concetto di perceptrone nel riconoscimento automatico di forme. In particolare, ha costruito una funzione peso per perceptroni per riconoscere siti di legame nelle sequenze di nucleotidi dell'RNA messaggero. Egli usa il teorema di convergenza dei perceptroni per portare le prestazioni del suo perceptrone a un livello ottimale. I risultati sono stati molto incoraggianti.

Qualsiasi perceptrone a finestra che includa una configurazione a finestra tutta bianca o una tutta nera è un buon perceptrone. Lowell Hill di Venice, California, lo ha notato e si è chiesto se una retina tutta bianca o tutta nera costituisca un'immagine legittima. La risposta dipende dalla configurazione. Nel caso del perceptrone a rettangoli multipli, sembra ragionevole considerare una retina tutta nera come un grande rettangolo.

In una fruttuosa incursione nel piccolo progetto di ricerca da me suggerito, Constantine Roussos del Lynchburg College Computer Center di Lynchburg, Virginia, ha deciso di escludere i perceptroni a finestra con configurazioni a finestra tutte bianche o tutte nere. Tra i risultati che ha raggiunto vi è una caratterizzazione dei buoni perceptroni (quelli che riconoscono almeno una configurazione). La caratterizzazione utilizza relazioni di traslazione tra le configurazioni a finestra sull'elenco del perceptrone. Se si sposta una configurazione a finestra di una singola unità in una qualsiasi delle quattro direzioni principali, si deve ottenere un'altra configurazione a finestra sull'elenco. Roussos ha poi concentrato la propria attenzione sui perceptroni a finestra minimali, quelli con un elenco che non può essere ulteriormente ridotto senza distruggere la «bontà» del perceptrone. Perceptroni di questo genere sono mattoni per la costruzione dell'insieme di tutti i buoni perceptroni. Roussos ha scritto un programma per calcolatore che scopre tutti i perceptroni a finestra minimali con elenchi di dimensioni varianti dall'ordine 2 all'ordine 5. Non esiste alcun perceptrone a finestra minimale di ordine 6.

John M. Evans di Hartford, Connecticut, ritiene che i fallimenti del perceptrone siano da addebitare alla restrizione derivante dalla gerarchia a due livelli, con demoni locali che riferiscono a un unico demone capo. Introducendo una sorta di «quadri intermedi» di demoni, Evans supera le limitazioni di connessione scoperte da Minsky e Papert. Gli stessi demoni di basso livello costituiscono una specie di retina con bianchi e neri a seconda che un particolare demone riferisca o non riferisca. Un secondo livello di demoni osserva la configurazione creata dai demoni di basso livello e riferisce la presenza di sottoconfigurazioni al demone capo. Un congegno a tre livelli è in grado di distinguere quali delle quattro configurazioni verificate sono connesse e quali no.